

# PROPAGATION OF CHAOS FOR THE KELLER-SEGEL EQUATION OVER BOUNDED DOMAINS

RAZVAN C. FETECAU, HUI HUANG, AND WEIRAN SUN

ABSTRACT. In this paper we rigorously justify the propagation of chaos for the parabolic-elliptic Keller-Segel equation over bounded convex domains. The boundary condition under consideration is the no-flux condition. As intermediate steps, we establish the well-posedness of the associated stochastic equation as well as the well-posedness of the Keller-Segel equation for bounded weak solutions.

**Keywords:** Coupling method, chemotaxis, Newtonian potential, mean-field limit, uniqueness, Wasserstein metric.

## 1. INTRODUCTION

In biology, *chemotaxis* describes the phenomenon that cells and organisms direct their movement in response to chemical gradients. This mechanism enables bacteria such as *Escherichia coli* (*E. coli*) to swim toward the highest concentration of food or flee from poisons. It also facilitates the movement of sperm towards the egg during fertilization. Mathematically, one of the most classical models for studying chemotaxis is the Keller-Segel (KS)-type equations, introduced in [35] to describe aggregation of the slime mold amoebae. There are many theoretical results regarding the well-posedness and mean-field limits of Keller-Segel-type equations when posed over the whole domain  $\mathbb{R}^d$  (see for example [3, 6, 18, 21, 27, 31, 55]).

In this paper, we aim at justifying the mean-field limit leading to the parabolic-elliptic Keller-Segel model in a bounded convex domain  $D$ . Specifically, let  $\rho_t$  be the macroscopic density and  $F$  be the interaction force among particles. Then the KS equation under investigation has the form

$$\begin{cases} \partial_t \rho_t = \nu \Delta \rho_t - \nabla \cdot [\rho_t F * \rho_t], & x \in D, \quad t > 0, \\ \rho_t(x)|_{t=0} = \rho_0(x), & x \in D, \\ \langle \nu \nabla \rho_t - \rho_t F * \rho_t, n \rangle = 0, & x \in \partial D, \end{cases} \quad (1.1)$$

where  $\langle \cdot, \cdot \rangle$  denotes the dot product in  $\mathbb{R}^d$  ( $d \geq 2$ ),  $n$  is the normal direction to  $\partial D$ , and the convolution is defined as

$$F * \rho_t(x) = \int_D F(x-y) \rho_t(y) dy, \quad \text{for any } x \in D.$$

The particular  $F$  considered in this paper is the negative gradient of the Newtonian potential given by

$$F(x) = -\nabla \Phi(x) = -\frac{C_* x}{|x|^d}, \quad \forall x \in \mathbb{R}^d \setminus \{0\}, \quad d \geq 2 \quad (1.2)$$

with  $C_* = \frac{\Gamma(d/2)}{2\pi^{d/2}}$  and

$$\Phi(x) = \begin{cases} \frac{1}{2\pi} \ln |x|, & \text{if } d = 2, \\ -\frac{C_d}{|x|^{d-2}}, & \text{if } d \geq 3. \end{cases} \quad (1.3)$$

Here,  $C_d = \frac{1}{d(d-2)\alpha_d}$  with  $\alpha_d = \frac{\pi^{d/2}}{\Gamma(d/2+1)}$  denoting the volume of  $d$ -dimensional unit ball. Equation (1.1) is also commonly written as a system of equations for the density  $\rho_t$  and the concentration  $c = -\Phi * \rho_t$  of the chemo-attractant. The main mechanism driving the solutions of (1.1) is the competition between the

diffusion and the nonlocal aggregation. We refer the reader to the review paper [23] or Chapter 5 in [41] for more discussions of this model.

The main purpose of this paper is to justify the propagation of chaos for the KS system (1.1). To do this, we model the dynamics of the particles by using equations on three levels: a continuity equation on the macroscopic level (1.1), a stochastic equation on the mesoscopic level and a stochastic interacting particle system on the microscopic level. The methodology in this paper is to approach the PDE model (1.1) from its underlying microscopic stochastic particle system. Specifically, let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  be a probability space endowed with a filtration  $(\mathcal{F}_t)_{t \geq 0}$  and  $\{B_t^i\}_{i=1}^N$  be  $N$  independent  $d$ -dimensional Brownian motions on this probability space. The system of interacting particles underlying the KS equation (1.1) is given by:

$$\begin{cases} X_t^i = X_0^i + \frac{1}{N-1} \sum_{j \neq i}^N \int_0^t F(X_s^i - X_s^j) ds + \sqrt{2\nu} B_t^i - R_t^i, & i = 1, \dots, N, \quad t > 0, \\ R_t^i = \int_0^t n(X_s^i) d|R^i|_s, \quad |R^i|_t = \int_0^t \mathbf{1}_{\partial D}(X_s^i) d|R^i|_s, & X_t^i|_{t=0} = X_0^i, \end{cases} \quad (1.4)$$

where  $\{X_t^i\}_{i=1}^N$  are the trajectories of  $N$  particles and  $F$  given by (1.2) models pairwise interaction between the particles. We assume that the initial data  $\{X_0^i\}_{i=1}^N$  are independently, identically distributed (i.i.d.) random variables with a common probability density function  $\rho_0(x)$  and  $\sqrt{2\nu}$  is a constant. Also,  $n(x)$  denotes the outward normal to  $\partial D$  at the point  $x \in \partial D$  and  $R_t^i$  is a reflecting process associated to  $X_t^i$  with a bounded total variation. Moreover,  $|R^i|_t$  is the total variation of  $R_t^i$  on  $[0, t]$ , namely

$$|R^i|_t = \sup \sum_k |R_{t_k}^i - R_{t_{k-1}}^i|, \quad (1.5)$$

where the supremum is taken over all partitions such that  $0 = t_0 < t_1 < \dots < t_n = t$ .

Solving the stochastic differential equation (SDE) (1.4) is known as the Skorokhod problem. Skorokhod introduced the problem in 1961 [43, 44], where he considered a SDE with a reflecting diffusion process on  $D = (0, \infty)$ . The multi-dimensional version of Skorokhod's problem was solved by Tanaka [50], where the domain  $D$  was assumed to be convex. In [37] Lions and Sznitman extended the result to a general domain satisfying certain admissibility conditions, where the admissibility roughly means that the domain can be approximated by smooth domains in a certain sense. Later, Saisho [42] removed the admissibility condition of the domain by applying the techniques used in [50].

The analysis of the limiting process of interacting particle systems as  $N \rightarrow \infty$  is usually called the *mean-field limit*. There is extensive literature [11, 28, 30, 46] on mean-field limits for various models when the spatial domain is the whole space  $\mathbb{R}^d$ . For second order systems, mean-field limits with globally Lipschitz forces in the whole space  $\mathbb{R}^d$  were obtained by Braun and Hepp [9] and Dobrushin [17]. Later, in [8], these results were extended to particle systems with locally Lipschitz interacting forces under certain moment control assumptions. The whole space case with a singular force was treated in [7, 10, 29, 36]. In particular, the mean-field limits for the KS system (1.1) in the whole space were shown in [21, 25–27, 55].

Meanwhile, there is much less work on mean-field limits for systems over bounded domains. In [48], the author derived rigorously the mean-field limit for particle systems with reflecting boundary conditions and bounded Lipschitz interaction forces. In Carrillo *et al.* [13], the authors considered a model in general domains (which is not necessarily convex) with a  $\lambda$ -geodesically convex potential. Very recently Choi and Salem [15] proved the mean-field limit for the case where the particle system has a particular type of bounded discontinuous interacting force.

The intermediate model between the microscopic model (1.4) and the macroscopic equation (1.1) is the following stochastic equation for the mean-field self-consistent stochastic process  $(Y_t)_{t \geq 0}$ :

$$\begin{cases} Y_t = Y_0 + \int_0^t \int_D F(Y_s - y) \rho_s(y) dy ds + \sqrt{2\nu} B_t - \tilde{R}_t, & t > 0, \\ \tilde{R}_t = \int_0^t n(Y_s) d|\tilde{R}|_s, \quad |\tilde{R}|_t = \int_0^t \mathbf{1}_{\partial D}(Y_s) d|\tilde{R}|_s, \end{cases} \quad (1.6)$$

where  $(\rho_t)_{t \geq 0}$  is the marginal density of  $(Y_t)_{t \geq 0}$  for any  $t \geq 0$  and  $Y_0$  has density  $\rho_0(x)$ . The solution of (1.6) is considered in the strong sense when the Brownian motion  $B_t$  is prescribed. By Itô's formula [40], it can be shown that the probability density function  $\rho_t$  of  $Y_t$  satisfies the KS equation (1.1) in the weak sense (see Definition 2.2 below).

The goal of the present manuscript is to justify the mean-field limit from the particle system (1.4) with the Newtonian force (1.2) to the mean-field KS equation (1.1) (see Theorem 2.7). More precisely, we show an explicit convergence rate from the weak solution of the mean-field equation (1.1) to the empirical measure associated with the particle system (1.4). The main idea is to introduce an approximate mean-field dynamics  $\{Y_t^i\}_{i=1}^N$  satisfying (1.6) and show that it well approximates the interacting particles  $\{X_t^i\}_{i=1}^N$  with a regularized  $F$  for large  $N$ 's. Our main contribution is to extend the propagation of chaos posed on the whole space with a singular force [21, 25–27, 55] or posed on bounded domains with bounded forces [15] to bounded convex domains with a singular force. We also show a universal upper bound of the expectation of the collision time in  $\mathbb{R}^2$  (see Theorem 5.1).

Before proving the propagation of chaos, we need to establish the well-posedness of solutions for the KS equation (1.1) and the stochastic equation (1.6). There is a huge literature on the well-posedness of the KS equation (see for example [3, 6, 18, 31] for the existence and to [14, 19, 34, 38, 47] for the uniqueness). There are also results for variants of the KS equation set on bounded domains with no-flux boundary conditions (see for example [1, 2, 4, 5, 16, 31, 39, 45, 45, 51, 52, 54]). In this paper, we use a coupling method to justify the well-posedness of equations (1.1) and (1.6). This coupling method has been used to show the well-posedness for the KS equation in the whole space in [24, 55]. Its main idea is to make use of the link between the KS equation (1.1) and the stochastic equation (1.6) driven by a Brownian motion. Our main contribution in this part is to extend such coupling method from  $\mathbb{R}^d$  to the bounded domain case.

The layout of the paper is as follows: in Section 2 we explain in detail the settings of our equations and the precise statements of the main theorems. In Section 3, we prove some technical lemmas that are used in the proofs. In Section 4, we establish existence, uniqueness, and stability of the KS equation (1.1) and the stochastic equation (1.6). Finally, we show the finite-time collision and propagation of chaos in Section 5.

## 2. SETTINGS AND MAIN RESULTS

In this section we explain the basic setting and main results in this paper. Our main approach in investigating the mean-field limit with a singular potential is to apply a regularization mechanism [25, 26]. To this end, let  $J(x)$  be a blob function such that

$$J \in C^\infty(\mathbb{R}^d), \quad J \geq 0, \quad \text{supp } J(x) \subset B(0, 1), \quad \int_{B(0,1)} J(x) dx = 1.$$

For any  $\varepsilon > 0$ , define  $J_\varepsilon(x) = \frac{1}{\varepsilon^d} J(\frac{x}{\varepsilon})$  and

$$F_\varepsilon(x) = J_\varepsilon * F(x). \tag{2.1}$$

The regularized version of (1.4) has the form

$$\begin{cases} X_t^{i,\varepsilon} = X_0^i + \frac{1}{N-1} \sum_{j \neq i}^N \int_0^t F_\varepsilon(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) ds + \sqrt{2\nu} B_t^{i,\varepsilon} - R_t^{i,\varepsilon}, & i = 1, \dots, N, \quad t > 0, \\ R_t^{i,\varepsilon} = \int_0^t n(X_s^{i,\varepsilon}) d|R^{i,\varepsilon}|_s, & |R^{i,\varepsilon}|_t = \int_0^t \mathbf{1}_{\partial D}(X_s^{i,\varepsilon}) d|R^{i,\varepsilon}|_s, \end{cases} \tag{2.2}$$

Here we assume that the initial data  $\{X_0^i\}_{i=1}^N$  are i.i.d. random variables with a common probability density function  $\rho_0(x)$ .

The global well-posedness of (2.2) is shown in [15, Theroem 1.1], which states

**Theorem 2.1** ([15]). *Let  $N \geq 2$ . For any initial data  $(X_0^i)_{i=1, \dots, N} \in \overline{D}^N$  and for any  $T > 0$ , there exists at least one weak solution  $(X_t^{i, \varepsilon}, R_t^{i, \varepsilon}, B_t^{i, \varepsilon})_{i=1, \dots, N}$  to system (2.2) on the time interval  $[0, T]$ .*

*Remark 2.1.* Note that the regularized system (2.2) is the same as the original system (1.4) before the collision time  $\tau$  defined in Theorem 5.1.

Next, we give the definition of weak solutions to the KS equation (1.1).

*Definition 2.2.* Assume  $D \subset \mathbb{R}^d$  is a bounded convex domain with boundary  $\partial D \in C^1$ . Let  $T > 0$  and  $0 \leq \rho_0(x) \in L^\infty(D)$ . A probability density function  $\rho_t(x) \in L^\infty(0, T; L^\infty(D))$  is a weak solution to (1.1) if

$$\begin{aligned} \int_D \rho_t(x) \varphi(x) dx &= \int_D \rho_0(x) \phi(x) dx + \nu \int_0^t \int_D \Delta \phi(x) \rho_s(x) dx ds \\ &\quad + \int_0^t \int_D \rho_s(x) F * \rho_s(x) \cdot \nabla \varphi(x) dx ds, \end{aligned} \quad (2.3)$$

for any test function  $\phi \in C^\infty(\overline{D})$  satisfying  $\langle \nabla \phi, n \rangle = 0$  on  $\partial D$ .

In our first main theorems we establish the well-posedness of (1.6) and (1.1) using the coupling method.

**Theorem 2.3.** (*Existence*) *Let  $Y_0$  be a random variable with density  $\rho_0 \in L^\infty(D)$ . Suppose the singular interaction force  $F$  is given by (1.2). Then there exists  $T > 0$  such that equation (1.6) has a strong solution  $(Y_t, \tilde{R}_t)$  and equation (1.1) with initial data  $\rho_0$  has a weak solution  $\rho_t \in L^\infty(0, T; L^\infty(D))$ . Moreover,  $\rho_t$  is the density of  $Y_t$  up to time  $T$ .*

One main tool we use in establishing the above well-posedness result is the Wasserstein distance. For the convenience of the reader, we give a brief introduction on the topology of the  $p$ -Wasserstein space. Consider the space

$$\mathcal{P}_p(\mathbb{R}^d) = \left\{ f \mid f \text{ is a probability measure on } \mathbb{R}^d \text{ and } \int_{\mathbb{R}^d} |x|^p df(x) < \infty \right\}. \quad (2.4)$$

The  $p$ -Wasserstein distance in  $\mathcal{P}_p(\mathbb{R}^d)$  is defined by

$$W_p(f, g) = \left( \inf_{\pi \in \Lambda(f, g)} \left\{ \int_{\mathbb{R}^{2d}} |x - y|^p d\pi(x, y) \right\} \right)^{1/p} = \left( \inf_{X \sim f, Y \sim g} \left\{ \mathbb{E}[|X - Y|^p] \right\} \right)^{1/p}, \quad 1 \leq p < \infty \quad (2.5)$$

and

$$W_\infty(f, g) := \inf_{X \sim f, Y \sim g} \mathbb{P}\text{-ess sup } |X - Y| = \inf_{\pi \in \Lambda(f, g)} \pi\text{-ess sup}_{(x, y) \in D \times D} |x - y|, \quad (2.6)$$

where  $\Lambda(f, g)$  is the set of joint probability measures on  $\mathbb{R}^d \times \mathbb{R}^d$  with marginals  $f$  and  $g$  respectively and  $(X, Y)$  are all possible couples of random variables with  $f$  and  $g$  as their distributions respectively. As a slight abuse of notation, we also use  $W_p(\rho_1, \rho_2)$  to denote the Wasserstein distance of two measures  $f, g$  with densities  $\rho_1, \rho_2$  respectively. In [53, Theorem 6.18], it has been shown that for any  $1 \leq p < \infty$ , the space  $\mathcal{P}_p(\mathbb{R}^d)$  endowed with the metric  $W_p$  is a complete metric space. Using the Wasserstein distance, we will establish the Dobrushin's type stability for (1.6) and (1.1). The result is given by the following theorem.

**Theorem 2.4.** (*Stability and Uniqueness*) (1) *If  $\{Y_t^i\}_{i=1,2}$  are two strong solutions to (1.6) with initial data  $\{Y_0^i\}_{i=1,2}$  and densities  $\{\rho_t^i\}_{i=1,2}$  respectively. Then there exist some constants  $C_1, C_2$  depending only on  $D, T, \|\rho_t^1\|_{L^\infty(0, T; L^\infty(D))}$  and  $\|\rho_t^2\|_{L^\infty(0, T; L^\infty(D))}$  such that*

$$\sup_{t \in [0, T]} \mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2| \leq C_1 \max \left\{ \mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2|, (\mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2|)^{\exp(-C_2 T)} \right\}. \quad (2.7)$$

(2) Let  $\{\rho_t^i\}_{i=1,2} \in L^\infty(0, T; L^\infty(D))$  be two weak solutions to (1.1) with initial data  $\{\rho_0^i\}_{i=1,2}$  respectively. Then there exist some constants  $C_1, C_2$  depending only on  $D, T, \|\rho_t^1\|_{L^\infty(0, T; L^\infty(D))}$  and  $\|\rho_t^2\|_{L^\infty(0, T; L^\infty(D))}$  such that

$$\sup_{t \in [0, T]} W_\infty(\rho_t^1, \rho_t^2) \leq C_1 \max \left\{ W_\infty(\rho_0^1, \rho_0^2), (W_\infty(\rho_0^1, \rho_0^2))^{\exp(-C_2 T)} \right\}. \quad (2.8)$$

Uniqueness then follows from the stability estimates.

Our second main theorem is to show the propagation of chaos and justify the mean-field limit of the regularized particle system (2.2). First, we recall the definition of a chaotic sequence [32, 49]:

*Definition 2.5* ( $\rho$ -chaotic). Let  $E$  be a Polish space and  $\rho$  a probability measure on  $E$ . Let  $\{X^i\}_{i=1}^N$  be a sequence of exchangeable random variables. Then  $\{X^i\}_{i=1}^N$  is  $\rho$ -chaotic if

$$\mu^N := \frac{1}{N} \sum_{i=1}^N \delta_{X^i} \rightarrow \rho \text{ as measures when } N \rightarrow \infty, \quad (2.9)$$

that is, for any  $\phi \in C_b(D)$ ,

$$\int_D \phi d\mu^N \rightarrow \int_D \phi \rho dx \quad \text{in law.}$$

A sufficient condition for the sequence  $\{X^i\}_{i=1}^N$  to be  $\rho$ -chaotic is

**Proposition 2.6.** [15, Remark 1.4] *Let  $\rho \in \mathcal{P}_p(E)$  with  $p \geq 1$ . Let  $\mu^N$  be the empirical measure defined in (2.9). If*

$$\mathbb{E} [W_p(\mu^N, \rho)] \rightarrow 0 \text{ as } N \rightarrow \infty, \quad (2.10)$$

then the sequence  $\{X^i\}_{i=1}^N$  is  $\rho$ -chaotic.

The result of propagation of chaos states

**Theorem 2.7.** (*Propagation of chaos*) *Let  $\rho_t \in L^\infty(0, T; L^\infty(D))$  be a weak solution to equation (1.1) with initial data  $\rho_0 \in L^\infty(D)$ . Assume that  $\{X_0^i\}_{i=1}^N$  are  $N$  i.i.d. random variables with the common density  $\rho_0$  and  $\{X_t^{i, \varepsilon}\}_{i=1}^N$  satisfy the particle system (2.2) with the initial data  $\{X_0^i\}_{i=1}^N$ . Let  $\mu_t^{X, \varepsilon}$  be the associated empirical measure given by*

$$\mu_t^{X, \varepsilon} := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i, \varepsilon}}. \quad (2.11)$$

Then there exist a cut-off parameter  $\varepsilon \sim \log^{-\frac{1}{d}}(N)$  and constants  $C_1, C_2, N_0 > 0$  depending only on  $\|\rho_0\|_{L^\infty(D)}$ ,  $D$  and  $T$  such that

$$\sup_{t \in [0, T]} \mathbb{E} \left[ W_2(\mu_t^{X, \varepsilon}, \rho_t) \right] \leq C_1 [\omega(\log^{-\frac{1}{d}}(N))]^{\exp(-C_2 T)} \quad \text{for all } N \geq N_0, \quad (2.12)$$

where  $\omega(x)$  is defined in (3.2).

*Remark 2.2.* In the theorem above, we assume that the initial data  $\{X_0^i\}_{i=1}^N$  are i.i.d. random variables with the common density  $\rho_0$ . This is a stronger assumption than  $\{X_0^i\}_{i=1}^N$  being  $\rho_0$ -chaotic. Our theorem shows that under this stronger assumption on the initial data, the system remains chaotic for a finite time.

## 3. PRELIMINARIES

In this section we collect the technical lemmas that are used in the proofs of the main theorems. Throughout this paper, we will denote any generic constants as  $C$ , which may change from line to line. The notation  $\|f\|_{L^p(D)}$  represents the usual  $L^p(D)$ -norm for any  $1 \leq p \leq \infty$ . We may suppress the dependence on  $D$  when there is no confusion.

The first lemma contains some useful properties of the regularized force  $F_\varepsilon$ , which states

**Lemma 3.1.** [55] *Let  $F_\varepsilon$  be the regularized force defined in (2.1). Then*

- (a)  $F_\varepsilon(-x) = -F_\varepsilon(x)$ ,  $F_\varepsilon(x) = F(x)$  for any  $|x| \geq \varepsilon$ , and  $|F_\varepsilon(x)| \leq |F(x)|$  for any  $x \in \mathbb{R}^d$ .
- (b)  $|\partial^\beta F_\varepsilon(x)| \leq C_\beta \varepsilon^{1-d-|\beta|}$  for any  $x \in \mathbb{R}^d$  and  $|\beta| \geq 0$ .

An immediate consequence of Lemma 3.1 is

**Lemma 3.2.** *For any function  $\rho \in L^\infty(D)$ , there exists a universal constant  $C$  that only depends on  $D$  such that for all  $\varepsilon \geq \varepsilon' \geq 0$ ,*

- (a)  $\|F_\varepsilon * \rho\|_{L^\infty(D)} \leq C \|\rho\|_{L^\infty(D)}$ .
- (b)  $\|F_\varepsilon * \rho - F_{\varepsilon'} * \rho\|_{L^\infty(D)} \leq C\varepsilon \|\rho\|_{L^\infty(D)}$ .

*Proof.* (a). By Lemma 3.1 (a), we have

$$|F_\varepsilon * \rho| \leq \left( \int_D F_\varepsilon(x-y) dy \right) \|\rho\|_{L^\infty(D)} \leq \left( \int_D \frac{C_*}{|x-y|^{d-1}} dy \right) \|\rho\|_{L^\infty(D)} \leq C \|\rho\|_{L^\infty(D)}.$$

(b). By Lemma 3.1 (a) again, we have

$$\begin{aligned} |F_\varepsilon * \rho(x) - F_{\varepsilon'} * \rho(x)| &\leq \int_D |F_\varepsilon(x-y) - F_{\varepsilon'}(x-y)| \rho(y) dy \\ &\leq C \|\rho\|_{L^\infty(D)} \int_{\{y: |x-y| \leq \varepsilon\}} \frac{dy}{|x-y|^{d-1}} \leq C \varepsilon \|\rho\|_{L^\infty(D)}, \end{aligned}$$

where  $C$  is a generic constant. □

Using the above two lemmas, we can prove a crucial quasi-log-Lipschitz estimate for  $F_\varepsilon$ .

**Lemma 3.3** (Quasi-log-Lipschitz). *Let  $X_1, X_2$  be two random variables with densities  $\rho_1, \rho_2 \in L^\infty(D)$  respectively ( $X_1$  and  $X_2$  may not be independent). Suppose  $r_0$  is large enough such that  $D \subseteq B(0, r_0)$ . Then for any  $0 \leq \varepsilon' \leq \varepsilon < 1$ , there exists a constant  $C$  depending only on  $D$ ,  $\|\rho_1\|_{L^\infty}$  and  $\|\rho_2\|_{L^\infty}$  such that*

$$|F_\varepsilon * \rho_1(X_1) - F_{\varepsilon'} * \rho_2(X_2)| \leq C\omega(\varepsilon) + C\omega(\mathbb{P}\text{-ess sup } |X_1 - X_2|), \quad (3.1)$$

where

$$\omega(r) = \begin{cases} 1, & \text{if } r \geq 1, \\ r(1 - \log r), & \text{if } 0 < r < 1. \end{cases} \quad (3.2)$$

*Proof.* Splitting the difference into two parts such that

$$\begin{aligned} F_\varepsilon * \rho_1(X_1) - F_{\varepsilon'} * \rho_2(X_2) &= (F_\varepsilon * \rho_1(X_1) - F_\varepsilon * \rho_2(X_2)) + (F_\varepsilon * \rho_2(X_2) - F_{\varepsilon'} * \rho_2(X_2)) \\ &=: I_1 + I_2, \end{aligned} \quad (3.3)$$

we will derive the desired bound by bounding  $I_1, I_2$  separately. The bound of  $I_2$  is readily obtained by Lemma 3.2 (b), which shows that there exists a constant  $C$  depending only on  $\|\rho_2\|_{L^\infty}$  such that

$$|I_2| \leq C\varepsilon. \quad (3.4)$$

The estimate of  $I_1$  can be obtained by following the approach in [33, Lemma 1.4]. To this end, we introduce the joint distribution of  $X_1$  and  $X_2$  as  $\pi := \mathcal{L}(X_1, X_2)$ . Then

$$|I_1| = \left| \iint_{D \times D} F_\varepsilon(X_1 - x_1) - F_\varepsilon(X_2 - x_2) \pi(\mathrm{d}x_1, \mathrm{d}x_2) \right|.$$

Let  $\ell$  be the constant defined by

$$\ell := \mathbb{P}\text{-ess sup } |X_1 - X_2| + \pi\text{-ess sup}_{(x,y) \in D \times D} |x - y|,$$

where

$$\mathbb{P}\text{-ess sup } |X_1 - X_2| := \inf \{ \lambda \geq 0 : \mathbb{P}(|X_1 - X_2| > \lambda) = 0 \},$$

and

$$\pi\text{-ess sup}_{(x,y) \in D \times D} |x - y| := \inf \{ \lambda \geq 0 : \pi(\{(x_1, x_2) \in D \times D : |x_1 - x_2| > \lambda\}) = 0 \}.$$

Note that by their definitions, we have

$$\mathbb{P}\text{-ess sup } |X_1 - X_2| = \pi\text{-ess sup}_{(x,y) \in D \times D} |x - y|.$$

Splitting  $|I_1|$  into two parts, we have

$$\begin{aligned} & \left| \iint_{D \times D} F_\varepsilon(X_1 - x_1) - F_\varepsilon(X_2 - x_2) \pi(\mathrm{d}x_1, \mathrm{d}x_2) \right| \\ & \leq \left| \iint_{(\{|X_1 - x_1| \leq 2\ell\} \cap D) \times D} F_\varepsilon(X_1 - x_1) - F_\varepsilon(X_2 - x_2) \pi(\mathrm{d}x_1, \mathrm{d}x_2) \right| \\ & \quad + \left| \iint_{(\{|X_1 - x_1| > 2\ell\} \cap D) \times D} F_\varepsilon(X_1 - x_1) - F_\varepsilon(X_2 - x_2) \pi(\mathrm{d}x_1, \mathrm{d}x_2) \right| \\ & =: I_{11} + I_{12}. \end{aligned}$$

Note that over the integration domain of  $I_{11}$ , we have

$$\begin{aligned} |X_2 - x_2| & \leq |X_1 - x_1| + |X_2 - X_1 + x_1 - x_2| \\ & \leq 2\ell + \mathbb{P}\text{-ess sup } |X_1 - X_2| + \pi\text{-ess sup}_{(x,y) \in D \times D} |x - y| = 3\ell. \end{aligned}$$

Therefore,  $I_{11}$  satisfies

$$\begin{aligned} I_{11} & \leq \int_{|X_1 - x_1| \leq 2\ell} \frac{C}{|X_1 - x_1|^{d-1}} \rho_1(x_1) \mathrm{d}x_1 + \int_{|X_2 - x_2| \leq 3\ell} \frac{C}{|X_2 - x_2|^{d-1}} \rho_2(x_2) \mathrm{d}x_2 \\ & \leq C(\|\rho_1\|_\infty + \|\rho_2\|_{L^\infty(D)})\ell, \end{aligned} \tag{3.5}$$

where we have applied the bound  $|F_\varepsilon(x)| \leq |F(x)|$ .

To bound  $|I_{12}|$ , we separate the two cases where  $2\ell > 2\varepsilon$  and  $2\ell \leq 2\varepsilon$ . First, if  $2\ell > 2\varepsilon$ , then  $|X_1 - x_1| > 2\ell > 2\varepsilon$ . By the definition of  $\ell$  we have

$$\ell \geq |X_1 - x_1| - |X_2 - x_2| > 2\ell - |X_2 - x_2|,$$

which implies  $|X_2 - x_2| > \ell > \varepsilon$ . Therefore, if  $2\ell > \varepsilon$ , then

$$F_\varepsilon(X_1 - x_1) = F(X_1 - x_1), \quad F_\varepsilon(X_2 - x_2) = F(X_2 - x_2),$$

and

$$F_\varepsilon(X_1 - x_1) - F_\varepsilon(X_2 - x_2) \leq \frac{C|X_2 - X_1 + x_1 - x_2|}{|X_1 - x_1 + \xi(X_2 - X_1 + x_1 - x_2)|^d} \quad \text{for some } \xi \in [0, 1]$$

Note that

$$\begin{aligned} |X_1 - x_1 + \xi(X_2 - X_1 + x_1 - x_2)| &\geq |X_1 - x_1| - |X_2 - X_1 + x_1 - x_2| \\ &\geq |X_1 - x_1| - \ell \geq |X_1 - x_1| - \frac{1}{2}|X_1 - x_1| = \frac{1}{2}|X_1 - x_1|. \end{aligned}$$

Therefore,

$$\begin{aligned} |I_{12}| &\leq \iint_{\{2\ell < |X_1 - x_1| \leq r_0\} \times D} |X_2 - X_1 + x_1 - x_2| \frac{C}{|X_1 - x_1 + \xi(X_2 - X_1 + x_1 - x_2)|^d} \pi(\mathrm{d}x_1, \mathrm{d}x_2) \\ &\leq C\ell \int_{2\ell < |X_1 - x_1| \leq r_0} \frac{2^d}{|X_1 - x_1|^d} \rho_1(x_1) \mathrm{d}x_1 \leq C \|\rho_1\|_{L^\infty} \ell \log\left(\frac{r_0}{2\ell}\right), \quad \text{if } 2\ell > 2\varepsilon. \end{aligned} \quad (3.6)$$

Thus it follows from (3.5) and (3.6) that

$$|I_1| \leq C\ell(1 - \log(\ell)), \quad \text{if } \varepsilon < \ell < 1. \quad (3.7)$$

In the case where  $2\ell \leq 2\varepsilon$ , we further divide  $I_{12}$  as

$$\begin{aligned} I_{12} &\leq \left| \iint_{(\{|X_1 - x_1| > 2\varepsilon\} \cap D) \times D} F_\varepsilon(X_1 - x_1) - F_\varepsilon(X_2 - x_2) \pi(\mathrm{d}x_1, \mathrm{d}x_2) \right| \\ &\quad + \iint_{(\{2\ell \leq |X_1 - x_1| \leq 2\varepsilon\} \cap D) \times D} |F_\varepsilon(X_1 - x_1) - F_\varepsilon(X_2 - x_2)| \pi(\mathrm{d}x_1, \mathrm{d}x_2) := I_{12}^{(1)} + I_{12}^{(2)}. \end{aligned}$$

The first term  $I_{12}^{(1)}$  is bounded in a similar as in the case  $2\ell > 2\varepsilon$ , which gives

$$I_{12}^{(1)} \leq C \|\rho_1\|_{L^\infty} \varepsilon \log\left(\frac{r_0}{2\varepsilon}\right). \quad (3.8)$$

By Lemma 3.1 (b), the second term  $I_{12}^{(2)}$  satisfies

$$\begin{aligned} I_{12}^{(2)} &\leq \iint_{(|X_1 - x_1| \leq 2\varepsilon) \cap D \times D} |X_1 - X_2 - x_1 + x_2| \varepsilon^{-d} \pi(\mathrm{d}x_1, \mathrm{d}x_2) \\ &\leq C\ell \varepsilon^{-d} \int_{|X_1 - x_1| \leq 2\varepsilon} \rho(x_1) \mathrm{d}x_1 \leq C\ell. \end{aligned} \quad (3.9)$$

Combining (3.8) and (3.9), we have

$$I_{12} \leq C\omega(\varepsilon), \quad \text{if } 2\ell \leq 2\varepsilon. \quad (3.10)$$

Collecting estimates (3.5) and (3.10), we derive that

$$|I_1| \leq C\omega(\varepsilon), \quad \text{if } 0 < \ell \leq \varepsilon.$$

Combing this with (3.7), we concludes that

$$|I_1| \leq C\omega(\varepsilon) + C\omega(\ell), \quad \text{if } 0 < \ell < 1.$$

where  $C$  depends only on  $D$ ,  $\|\rho_1\|_{L^\infty(D)}$  and  $\|\rho_2\|_{L^\infty(D)}$ . For  $\ell \geq 1$ , it follows from Lemma 3.2 (i) that  $|I_1| \leq C$ . We thereby conclude our proof.  $\square$

There are two Gronwall-type inequalities used in the later proofs. The first one is

**Lemma 3.4.** [15, Lemma A.1] *Let  $f$  be a nonnegative scalar function satisfying*

$$f(t) \leq f_0 \exp\left(C \int_0^t f(s) \mathrm{d}s\right), \quad t \geq 0.$$

Then

$$f(t) \leq \frac{f_0}{1 - Cf_0 t}, \quad t \geq 0,$$

where  $C$  is a positive constant.



The second Gronwall-type inequality states

**Lemma 3.5.** *Let  $\{f_\varepsilon(t)\}_{\varepsilon>0}$  be a family of nonnegative continuous functions satisfying*

$$f_\varepsilon^2(t) \leq C \int_0^t f_\varepsilon(s) \omega(f_\varepsilon(s)) ds + C\omega(\varepsilon)T \quad \text{for all } t \in [0, T], \quad (3.11)$$

where  $C > 0$  is a constant,  $T$  is independent of  $\varepsilon$ , and  $\omega(x)$  is defined in (3.2). Suppose  $\varepsilon$  is small enough such that  $C\omega(\varepsilon)T < 1$ . Then there exist two constants  $C_T$  and  $0 < \Lambda_0 < 1/2$  depending only on  $T$  and  $C$  such that if  $\omega(\varepsilon) \leq \Lambda_0$ , then

$$\sup_{t \in [0, T]} f_\varepsilon(t) \leq C_T \omega(\varepsilon)^{\frac{1}{2} e^{-CT}} < 1. \quad (3.12)$$

*Proof.* Let  $R_\varepsilon$  be the continuous function defined by

$$R_\varepsilon(t) := C\omega(\varepsilon)T + \int_0^t C f_\varepsilon(s) \omega(f_\varepsilon(s)) ds.$$

Since  $R_\varepsilon(0) < 1$  and  $R_\varepsilon(t)$  is continuous in time, there exists  $T^* > 0$  such that  $R_\varepsilon(t) < 1$  for all  $t \in [0, T^*]$ . We will show that such  $T^*$  can be extended to  $T$  by choosing  $\varepsilon$  small enough. Indeed, for  $\omega(\varepsilon) < 1/2$ , define

$$F(s) := \exp(1 - \exp(-Cs)) \omega(\varepsilon)^{\frac{1}{2} \exp(-Cs)} \quad \text{and} \quad T_\varepsilon := \frac{1}{C} \log(1 - \frac{1}{2} \log(\omega(\varepsilon))). \quad (3.13)$$

Then  $F$  is an increasing function and the particular choice of  $T_\varepsilon$  satisfies

$$F(T_\varepsilon) = 1, \quad T_\varepsilon \geq \frac{1}{C} \log\left(1 + \frac{1}{2} \log 2\right) > 0, \quad \omega(\varepsilon) < 1/2. \quad (3.14)$$

Now we show that  $0 < R_\varepsilon(t) < 1$  for all  $t \in [0, T_\varepsilon]$ . We prove this by contradiction. Assume there exists  $T_1 < T_\varepsilon$  such that

$$R_\varepsilon(t) < 1 \quad \text{for any } t \in [0, T_1] \quad \text{and} \quad R_\varepsilon(T_1) = 1.$$

Since  $\omega(x) = x(1 - \log x)$  on  $(0, 1]$  and it is nondecreasing, by (3.11) we have

$$\frac{dR_\varepsilon(t)}{dt} = C f_\varepsilon(t) \omega(f_\varepsilon(t)) \leq C(R_\varepsilon(t))^{1/2} \omega((R_\varepsilon(t))^{1/2}), \quad (3.15)$$

which gives

$$\frac{d}{dt} (R_\varepsilon(t))^{\frac{1}{2}} = \frac{1}{2} (R_\varepsilon(t))^{-\frac{1}{2}} \frac{dR_\varepsilon(t)}{dt} \leq \frac{C}{2} \omega\left((R_\varepsilon(t))^{1/2}\right) \leq C(R_\varepsilon(t))^{1/2} \left(1 - \log((R_\varepsilon(t))^{\frac{1}{2}})\right).$$

Solving this differential inequality on  $t \in [0, T_1]$ , we get

$$(R_\varepsilon(t))^{\frac{1}{2}} \leq e^{1 - (1 - \frac{1}{2} \log(C\omega(\varepsilon)T_1))e^{-CT_1}} \leq \exp(1 - \exp(-CT_1)) \omega(\varepsilon)^{\frac{1}{2} \exp(-CT_1)} = F(T_1). \quad (3.16)$$

By the definitions and properties of  $F, T_1, T_\varepsilon$ , we have

$$(R_\varepsilon(T_1))^{\frac{1}{2}} \leq F(T_1) < F(T_\varepsilon) = 1,$$

which contradicts the assumption  $R_\varepsilon(T_1) = 1$ . Hence we have proved that  $R_\varepsilon(t) < 1$  for all  $t \in [0, T_\varepsilon]$  with  $T_\varepsilon$  defined in (3.13). Notice that by the definition of  $T_\varepsilon$ , there exists  $\Lambda_0 < 1/2$  depending only on  $C$  and  $T$  such that  $T_\varepsilon > T$  for all  $\omega(\varepsilon) \leq \Lambda_0$ . For such  $\varepsilon$  we have

$$R_\varepsilon(t) < 1 \quad \text{for all } t \in [0, T].$$

Then repeating the derivation of (3.16), we obtain that

$$f_\varepsilon(t) \leq (R_\varepsilon(t))^{\frac{1}{2}} \leq F(T) \leq C_T \omega(\varepsilon)^{\frac{1}{2} e^{-CT}} \quad \text{for all } t \in [0, T].$$

□

## 4. WELL-POSEDNESS OF THE NONLINEAR SDE AND PDE

In this section we establish the well-posedness of the nonlinear SDE (1.6) and show that the marginal density of its solution is the unique weak solution to the KS equation (1.1).

**4.1. Regularized system.** The existence of solutions to (1.6) and (1.1) will be obtained through the convergence of approximate solutions to regularized versions of (1.6) and (1.1), where the singular force  $F$  is replaced by  $F_\varepsilon$  defined in (2.1).

First, we show the well-posedness of the regularized PDE

$$\begin{cases} \partial_t \rho_t^\varepsilon = \nu \Delta \rho_t^\varepsilon - \nabla \cdot [\rho_t^\varepsilon F_\varepsilon * \rho_t^\varepsilon], & x \in D, \quad t > 0, \\ \rho_t^\varepsilon(x)|_{t=0} = \rho_0(x), & x \in D, \\ \langle \nu \nabla \rho_t^\varepsilon - \rho_t^\varepsilon F_\varepsilon * \rho_t^\varepsilon, n \rangle = 0, & \text{on } \partial D. \end{cases} \quad (4.1)$$

**Proposition 4.1.** *Suppose  $\rho_0 \in L^\infty(D)$ . Then there exists  $T > 0$  independent of  $\varepsilon$  such that equation (4.1) has a weak solution  $\rho_t^\varepsilon \in L^\infty(0, T; L^\infty(D))$  with the bound*

$$\sup_{0 \leq t \leq T} \|\rho_t^\varepsilon\|_{L^\infty(D)} \leq 2 \|\rho_0\|_{L^\infty(D)}. \quad (4.2)$$

*Proof.* In order to show the existence of a weak solution to (4.1), first we consider the regularized equation

$$\begin{cases} \partial_t \rho_t^{\varepsilon, \delta} = \nu \Delta \rho_t^{\varepsilon, \delta} - \nabla \cdot [\rho_t^{\varepsilon, \delta} F_\varepsilon * \rho_t^{\varepsilon, \delta}], & x \in D, \quad t > 0, \\ \rho_t^{\varepsilon, \delta}(x)|_{t=0} = \rho_0^\delta(x), & x \in D, \\ \langle \nu \nabla \rho_t^{\varepsilon, \delta} - \rho_t^{\varepsilon, \delta} F_\varepsilon * \rho_t^{\varepsilon, \delta}, n \rangle = 0 & \text{on } \partial D, \end{cases} \quad (4.3)$$

where  $\rho_0^\delta \in C^\infty(D)$  satisfies

$$\|\rho_0^\delta\|_{L^\infty(D)} \leq \|\rho_0\|_{L^\infty(D)}, \quad \rho_0^\delta \rightarrow \rho_0 \quad \text{in } L^1 \text{ as } \delta \rightarrow 0.$$

Then for any fixed  $\varepsilon, \delta > 0$ , there exists  $T_1$  (which may depend on  $\varepsilon, \delta$ ) such that system (4.3) has a smooth nonnegative solution  $\rho_t^{\varepsilon, \delta} \in C([0, T_1]; L^\infty(D))$  c.f. [12]. In order to remove the dependence of  $T_1$  on  $\varepsilon, \delta$ , we show (a priori) that there exist  $T, C > 0$  which only depend on  $\|\rho_0\|_{L^\infty}$  and the domain  $D$  such that

$$\sup_{t \in [0, T]} \|\rho_t^{\varepsilon, \delta}\|_{L^\infty(D)} \leq C. \quad (4.4)$$

Indeed, for any  $p \geq 2$ , we have

$$\begin{aligned} \frac{d \|\rho_t^{\varepsilon, \delta}\|_p^p}{dt} &= p \int_D (\rho_t^{\varepsilon, \delta})^{p-1} (\nu \Delta \rho_t^{\varepsilon, \delta} - \nabla \cdot (\rho_t^{\varepsilon, \delta} F_\varepsilon * \rho_t^{\varepsilon, \delta})) dx \\ &= p\nu \int_{\partial D} (\rho_t^{\varepsilon, \delta})^{p-1} \nabla \rho_t^{\varepsilon, \delta} \cdot n \, dS(x) - p\nu \int_D \nabla \rho_t^{\varepsilon, \delta} \cdot \nabla ((\rho_t^{\varepsilon, \delta})^{p-1}) dx \\ &\quad - p \int_{\partial D} (\rho_t^{\varepsilon, \delta})^p F_\varepsilon * \rho_t^{\varepsilon, \delta} \cdot n \, dS(x) + p \int_D \rho_t^{\varepsilon, \delta} F_\varepsilon * \rho_t^{\varepsilon, \delta} \cdot \nabla ((\rho_t^{\varepsilon, \delta})^{p-1}) dx \\ &= -p(p-1)\nu \int_D (\rho_t^{\varepsilon, \delta})^{p-2} |\nabla \rho_t^{\varepsilon, \delta}|^2 dx + p \int_D \rho_t^{\varepsilon, \delta} F_\varepsilon * \rho_t^{\varepsilon, \delta} \cdot \nabla ((\rho_t^{\varepsilon, \delta})^{p-1}) dx. \end{aligned} \quad (4.5)$$

re-organize equation (4.5) as

$$\begin{aligned} \frac{d \|\rho_t^{\varepsilon, \delta}\|_p^p}{dt} + \frac{4\nu(p-1)}{p} \left\| \nabla ((\rho_t^{\varepsilon, \delta})^{\frac{p}{2}}) \right\|_2^2 &= (p-1) \int_D F_\varepsilon * \rho_t^{\varepsilon, \delta} \cdot \nabla ((\rho_t^{\varepsilon, \delta})^p) dx \\ &= 2(p-1) \int_D (\rho_t^{\varepsilon, \delta})^{\frac{p}{2}} F_\varepsilon * \rho_t^{\varepsilon, \delta} \cdot \nabla ((\rho_t^{\varepsilon, \delta})^{\frac{p}{2}}) dx. \end{aligned} \quad (4.6)$$

By (a) in Lemma 3.2, one has

$$\int_D (\rho_t^{\varepsilon,\delta})^{\frac{p}{2}} F_\varepsilon * \rho_t^{\varepsilon,\delta} \cdot \nabla((\rho_t^{\varepsilon,\delta})^{\frac{p}{2}}) dx \leq C \left\| \rho_t^{\varepsilon,\delta} \right\|_{L^\infty(D)} \left\| (\rho_t^{\varepsilon,\delta})^{\frac{p}{2}} \right\|_2 \left\| \nabla((\rho_t^{\varepsilon,\delta})^{\frac{p}{2}}) \right\|_2. \quad (4.7)$$

Applying Cauchy-Schwarz inequality in (4.6), we obtain

$$\frac{d \left\| \rho_t^{\varepsilon,\delta} \right\|_p^p}{dt} + \frac{4\nu(p-1)}{p} \left\| \nabla((\rho_t^{\varepsilon,\delta})^{\frac{p}{2}}) \right\|_2^2 \leq \frac{2\nu(p-1)}{p} \left\| \nabla((\rho_t^{\varepsilon,\delta})^{\frac{p}{2}}) \right\|_2^2 + Cp \left\| \rho_t^{\varepsilon,\delta} \right\|_{L^\infty(D)}^2 \left\| \rho_t^{\varepsilon,\delta} \right\|_p^p, \quad (4.8)$$

where  $C$  is independent of  $\varepsilon, \delta$ , and  $p$ . Gronwall's inequality then leads to

$$\left\| \rho_t^{\varepsilon,\delta} \right\|_p \leq \left\| \rho_0^\delta \right\|_p \exp \left( C \int_0^t \left\| \rho_s^\varepsilon \right\|_{L^\infty(D)}^2 ds \right) \quad \text{for any } p \geq 2. \quad (4.9)$$

Now we send  $p \rightarrow \infty$  to get

$$\left\| \rho_t^{\varepsilon,\delta} \right\|_{L^\infty(D)} \leq \left\| \rho_0^\delta \right\|_{L^\infty(D)} \exp \left( C \int_0^t \left\| \rho_s^\varepsilon \right\|_{L^\infty(D)}^2 ds \right). \quad (4.10)$$

Finally, applying Lemma 3.4 with  $f(t) = \left\| \rho_t^{\varepsilon,\delta} \right\|_{L^\infty(D)}^2$  yields that

$$\left\| \rho_t^{\varepsilon,\delta} \right\|_{L^\infty(D)}^2 \leq \frac{\left\| \rho_0^\delta \right\|_{L^\infty(D)}^2}{1 - 2C \left\| \rho_0^\delta \right\|_{L^\infty(D)}^2 t} \leq \frac{\left\| \rho_0 \right\|_{L^\infty(D)}^2}{1 - 2C \left\| \rho_0 \right\|_{L^\infty(D)}^2 t}. \quad (4.11)$$

Therefore, the uniform existence time  $T$  can be chose as

$$T = \frac{3}{8C \left\| \rho_0 \right\|_{L^\infty(D)}^2},$$

such that

$$\left\| \rho_t^{\varepsilon,\delta} \right\|_{L^\infty} \leq 2 \left\| \rho_0 \right\|_{L^\infty(D)} \quad \text{on } [0, T].$$

Therefore there exists a  $\rho_t^\varepsilon \in L^\infty(0, T; L^\infty(D))$  such that up to a subsequence,

$$\rho_t^{\varepsilon,\delta} \rightharpoonup^* \rho_t^\varepsilon \quad \text{in } L^\infty(0, T; L^\infty(D)). \quad (4.12)$$

Now we show that  $\rho_t^\varepsilon$  is a weak solution to (4.1). Notice that for any  $\phi \in C^\infty(\overline{D})$  satisfying  $\langle \nabla\phi(x), n(x) \rangle = 0$  on  $\partial D$ , we have

$$\begin{aligned} \int_D \rho_t^{\varepsilon,\delta}(x) \phi(x) dx &= \int_D \rho_0^\delta(x) \phi(x) dx + \nu \int_0^t \int_D \Delta\phi(x) \rho_s^{\varepsilon,\delta}(x) dx ds \\ &\quad + \int_0^t \int_D \rho_s^{\varepsilon,\delta}(x) F_\varepsilon * \rho_s^{\varepsilon,\delta}(x) \cdot \nabla\phi(x) dx ds. \end{aligned} \quad (4.13)$$

Then by taking a continuous version of the left-hand side and (4.12), we derive that

$$\begin{aligned} \int_D \rho_t^{\varepsilon,\delta}(x) \phi(x) dx &\rightarrow \int_D \rho_t^\varepsilon(x) \phi(x) dx, & \int_D \rho_0^\delta(x) \phi(x) dx &\rightarrow \int_D \rho_0(x) \phi(x) dx, \\ \nu \int_0^t \int_D \Delta\phi(x) \rho_s^{\varepsilon,\delta}(x) dx ds &\rightarrow \nu \int_0^t \int_D \Delta\phi(x) \rho_s^\varepsilon(x) dx ds \end{aligned}$$

as  $\delta \rightarrow 0$ . For the last term on the right-hand side of (4.13), we have

$$\begin{aligned}
& \left| \int_0^t \int_D \rho_s^{\varepsilon, \delta}(x) F_\varepsilon * \rho_s^{\varepsilon, \delta}(x) \cdot \nabla \phi(x) \, dx \, ds - \int_0^t \int_D \rho_s^\varepsilon(x) F_\varepsilon * \rho_s^\varepsilon(x) \cdot \nabla \phi(x) \, dx \, ds \right| \\
& \leq \left| \int_0^t \int_D \rho_s^{\varepsilon, \delta}(x) F_\varepsilon * \rho_s^{\varepsilon, \delta}(x) \cdot \nabla \phi(x) \, dx \, ds - \int_0^t \int_D \rho_s^{\varepsilon, \delta}(x) F_\varepsilon * \rho_s^\varepsilon(x) \cdot \nabla \phi(x) \, dx \, ds \right| \\
& \quad + \left| \int_0^t \int_D \rho_s^{\varepsilon, \delta}(x) F_\varepsilon * \rho_s^\varepsilon(x) \cdot \nabla \phi(x) \, dx \, ds - \int_0^t \int_D \rho_s^\varepsilon(x) F_\varepsilon * \rho_s^\varepsilon(x) \cdot \nabla \phi(x) \, dx \, ds \right| \\
& \leq C \int_0^t \int_D |F_\varepsilon * (\rho_s^{\varepsilon, \delta} - \rho_s^\varepsilon)(x)| \, ds \, dx + C \left| \int_0^t \int_D (\rho_s^{\varepsilon, \delta} - \rho_s^\varepsilon) \phi(x) \, dx \, ds \right|. \tag{4.14}
\end{aligned}$$

By (4.12), one has

$$\left| \int_0^t \int_D (\rho_s^{\varepsilon, \delta} - \rho_s^\varepsilon) \phi(x) \, dx \, ds \right| \rightarrow 0 \quad \text{as } \delta \rightarrow 0. \tag{4.15}$$

Notice that for any  $x \in D$ , one has

$$|F_\varepsilon * (\rho_s^{\varepsilon, \delta} - \rho_s^\varepsilon)(x)| \leq \int_D F_\varepsilon(x-y) (\rho_s^{\varepsilon, \delta}(y) - \rho_s^\varepsilon(y)) \, dy \rightarrow 0 \quad \text{as } \delta \rightarrow 0, \tag{4.16}$$

since  $F_\varepsilon \in L^1(D)$ . It follows from the dominated convergence theorem that

$$\int_0^t \int_D |F_\varepsilon * (\rho_s^{\varepsilon, \delta} - \rho_s^\varepsilon)(x)| \, ds \, dx \rightarrow 0 \quad \text{as } \delta \rightarrow 0. \tag{4.17}$$

Therefore,

$$\int_0^t \int_D \rho_s^{\varepsilon, \delta}(x) F_\varepsilon * \rho_s^{\varepsilon, \delta}(x) \cdot \nabla \phi(x) \, dx \, ds \rightarrow \int_0^t \int_D \rho_s^\varepsilon(x) F_\varepsilon * \rho_s^\varepsilon(x) \cdot \nabla \phi(x) \, dx \, ds \quad \text{as } \delta \rightarrow 0. \tag{4.18}$$

Combining the above limits, we have shown that  $\rho_t^\varepsilon$  is a weak solution to (4.1) in the sense that

$$\begin{aligned}
\int_D \rho_t^\varepsilon(x) \phi(x) \, dx &= \int_D \rho_0(x) \phi(x) \, dx + \nu \int_0^t \int_D \Delta \phi(x) \rho_s^\varepsilon(x) \, dx \, ds \\
&\quad + \int_0^t \int_D \rho_s^\varepsilon(x) F_\varepsilon * \rho_s^\varepsilon(x) \cdot \nabla \phi(x) \, dx \, ds.
\end{aligned}$$

with the bound in (4.2).  $\square$

Next, we show that for any weak solution  $\rho_t^\varepsilon \in L^\infty(0, T; L^\infty(D))$  to (4.1). One can construct a process  $Y_t^\varepsilon$  with  $\rho_t^\varepsilon$  being its density function, which satisfies regularized self-consistent SDE

$$\begin{cases} Y_t^\varepsilon = Y_0 + \int_0^t \int_D F_\varepsilon(Y_s^\varepsilon - y) \rho_s^\varepsilon(y) \, dy \, ds + \sqrt{2\nu} B_t - \tilde{R}_t^\varepsilon, & t > 0, \\ \tilde{R}_t^\varepsilon = \int_0^t n(Y_s^\varepsilon) \, d\tilde{R}^\varepsilon|_s, & |\tilde{R}^\varepsilon|_t = \int_0^t \mathbf{1}_{\partial D}(Y_s^\varepsilon) \, d|\tilde{R}^\varepsilon|_s, \end{cases} \tag{4.19}$$

where we require  $Y_t^\varepsilon$  possessing a density function  $\rho_t^\varepsilon(x)$ . The result states as follows

**Proposition 4.2.** *If  $\rho_t^\varepsilon \in L^\infty(0, T; L^\infty(D))$  is a weak solution to (4.1) with the initial data  $\rho_0 \in L^\infty(D)$ , then given any random variable  $Y_0$  with the density  $\rho_0$ , there is a process  $Y_t^\varepsilon$  with the marginal density  $\rho_t^\varepsilon$  satisfying (4.19) with the initial data  $Y_0$ .*

*Proof.* Consider the following linear PDE

$$\begin{cases} \partial_t \rho_t^\varepsilon = \nu \Delta \rho_t^\varepsilon - \nabla \cdot [\rho_t^\varepsilon V_\varepsilon[g_t]], & x \in D, \quad t > 0, \\ \rho_t(x)|_{t=0} = \rho_0(x), & x \in D \text{ and } \langle \nu \nabla \rho_t^\varepsilon - \rho_t V_\varepsilon[g_t], n \rangle = 0 \text{ on } \partial D, \end{cases} \tag{4.20}$$

where  $V_\varepsilon[g_t] = \int_D F_\varepsilon(x-y) g_t(y) \, dy$ ,  $g_t(x) \in L^\infty(0, T; L^\infty(D))$  is a given function. According to Lemma 3.2 (i), the term  $V_\varepsilon[g_t]$  is bounded, thus the linear PDE (4.20) has a unique weak solution in  $L^\infty(0, T; L^1(D))$ .

Next let  $V_\varepsilon[\rho_t^\varepsilon] = \int_D F_\varepsilon(x-y)\rho_t^\varepsilon(y)dy$ , where  $\rho_t^\varepsilon \in L^\infty(0, T; L^\infty(D))$  is a weak solution to (4.1). According to [48, Theorem 1.1], the following SDE

$$\begin{cases} Y_t^\varepsilon = Y_0 + \int_0^t V_\varepsilon[\rho_s^\varepsilon](Y_s^\varepsilon) ds + \sqrt{2\nu}B_t - \tilde{R}_t^\varepsilon, & t > 0, \\ \tilde{R}_t^\varepsilon = \int_0^t n(Y_s^\varepsilon) d|\tilde{R}^\varepsilon|_s, & |\tilde{R}^\varepsilon|_t = \int_0^t \mathbf{1}_{\partial D}(Y_s^\varepsilon) d|\tilde{R}^\varepsilon|_s, \end{cases} \quad (4.21)$$

has a strong solution  $Y_t^\varepsilon$  with density  $\bar{\rho}_t^\varepsilon$  since  $V_\varepsilon[\rho_t^\varepsilon]$  is smooth. For  $\phi \in C^\infty(\bar{D})$  satisfying  $\langle \nabla\phi(x), n(x) \rangle = 0$  for  $x \in \partial D$ , one applies Itô's formula:

$$\begin{aligned} \phi(Y_t^\varepsilon) &= \phi(Y_0) + \int_0^t \langle \nabla\phi(Y_s^\varepsilon), dY_s^\varepsilon \rangle + \frac{1}{2} \int_0^t (dY_s^\varepsilon)^T \nabla^2 \phi(Y_s^\varepsilon) dY_s^\varepsilon \\ &= \phi(Y_0) + \int_0^t \langle \nabla\phi(Y_s^\varepsilon), V_\varepsilon[\rho_s^\varepsilon](Y_s^\varepsilon) \rangle ds + \sqrt{2\nu} \int_0^t \langle \nabla\phi(Y_s^\varepsilon), dB_s \rangle \\ &\quad - \int_0^t \langle \nabla\phi(Y_s^\varepsilon), d\tilde{R}_s^\varepsilon \rangle + \nu \int_0^t \Delta\phi(Y_s^\varepsilon) ds. \end{aligned} \quad (4.22)$$

Using the boundary condition, one has

$$\int_0^t \langle \nabla\phi(Y_s^\varepsilon), d\tilde{R}_s^\varepsilon \rangle = \int_0^t \langle \nabla\phi(Y_s^\varepsilon), n(Y_s^\varepsilon) \rangle d|\tilde{R}^\varepsilon|_s = 0. \quad (4.23)$$

Taking expectation on (4.22), it follows from above that

$$\int_D \bar{\rho}_t^\varepsilon(x)\phi(x) dx = \int_D \rho_0(x)\phi(x) dx + \nu \int_0^t \int_D \Delta\phi(x)\bar{\rho}_s^\varepsilon(x) dx ds + \int_0^t \int_D \bar{\rho}_s^\varepsilon(x)V_\varepsilon[\rho_t^\varepsilon] \cdot \nabla\phi(x) dx ds.$$

This implies that  $\bar{\rho}_t^\varepsilon$  is a weak solution to (4.20) associated to  $V_\varepsilon[\rho_t^\varepsilon]$ . Notice that  $\rho_t^\varepsilon$  is also a weak solution to (4.20) associated to  $V_\varepsilon[\rho_t^\varepsilon]$ . By the uniqueness of weak solution to (4.20), we obtain  $\rho_t^\varepsilon = \bar{\rho}_t^\varepsilon$ . Hence we have proved that there exists a process  $Y_t^\varepsilon$  with density  $\rho_t^\varepsilon$  which is a strong solution to the regularized mean-field dynamics (4.19).  $\square$

**4.2. The original system.** Now we show the proof of Theorem 2.3, which asserts the existence of solutions to the SDE (1.6) and the KS equation (1.1).

*Proof of Theorem 2.3.* Note that the existence of weak solutions to the KS equation (1.1) follows from applying the Itô's formula on SDE (1.6). Thus we focus on proving the existence of a strong solution to the SDE (1.6), which is obtained by first constructing approximate solutions to the regularized SDE (4.19) and then passing to the limit.

To this end, we let  $\varepsilon > \varepsilon' > 0$  be arbitrary and  $F_\varepsilon, F_{\varepsilon'}$  be the corresponding regularized forces. Denote  $\rho_t^\varepsilon, \rho_t^{\varepsilon'}$  are two weak solutions to (4.1) with the same initial condition  $\rho_0 \in L^\infty(D)$ . Then there exists  $T > 0$  independent of  $\varepsilon, \varepsilon'$  such that  $\|\rho_t^\varepsilon\|_{L^\infty(D)}$  and  $\|\rho_t^{\varepsilon'}\|_{L^\infty(D)}$  are uniformly bounded in  $\varepsilon$  and  $\varepsilon'$  on  $[0, T]$ .

According to Proposition 4.2, we can construct two processes  $Y_t^\varepsilon, Y_t^{\varepsilon'}$  satisfying (4.19) with the same initial data  $Y_0$  and Brownian motion  $B_t$ , which have density functions  $\rho_t^\varepsilon$  and  $\rho_t^{\varepsilon'}$  respectively. We shall show that there exist constants  $C, C_T$ , and  $\varepsilon_0$  depending only on  $D, T$  and  $\|\rho_0\|_\infty$  such that if  $0 < \varepsilon' < \varepsilon < \varepsilon_0$ , then

$$\sup_{t \in [0, T]} \mathcal{W}_\infty(\rho_t^\varepsilon, \rho_t^{\varepsilon'}) \leq \sup_{t \in [0, T]} \mathbb{P}\text{-ess sup } |Y_t^\varepsilon - Y_t^{\varepsilon'}| \leq C_T \omega(\varepsilon)^{\frac{1}{2}e^{-C_T}}. \quad (4.24)$$

Indeed, by applying Itô's formula to  $Y_t^\varepsilon - Y_t^{\varepsilon'}$ , we have

$$\begin{aligned} |Y_t^\varepsilon - Y_t^{\varepsilon'}|^2 &= 2 \int_0^t \langle Y_s^\varepsilon - Y_s^{\varepsilon'}, F_\varepsilon * \rho_s^\varepsilon(Y_s^\varepsilon) - F_{\varepsilon'} * \rho_s^{\varepsilon'}(Y_s^{\varepsilon'}) \rangle ds \\ &\quad - 2 \int_0^t \langle Y_s^\varepsilon - Y_s^{\varepsilon'}, n(Y_s^\varepsilon) \rangle d|\tilde{R}^\varepsilon|_s - 2 \int_0^t \langle Y_s^{\varepsilon'} - Y_s^\varepsilon, n(Y_s^{\varepsilon'}) \rangle d|\tilde{R}^{\varepsilon'}|_s. \end{aligned}$$

By the convexity of the domain  $D$ , one has

$$(x - y) \cdot n(x) \geq 0 \quad \text{for any } x \in \partial D \text{ and } y \in \bar{D}.$$

Consequently,

$$\langle Y_s^\varepsilon - Y_s^{\varepsilon'}, n(Y_s^\varepsilon) \rangle \geq 0, \quad d|\tilde{R}^\varepsilon|_s \text{ almost surely}$$

and

$$\langle Y_s^{\varepsilon'} - Y_s^\varepsilon, n(Y_s^{\varepsilon'}) \rangle \geq 0, \quad d|\tilde{R}^{\varepsilon'}|_s \text{ almost surely}.$$

Therefore, by Lemma 3.3, we have

$$\begin{aligned} |Y_t^\varepsilon - Y_t^{\varepsilon'}|^2 &\leq 2 \int_0^t \langle Y_s^\varepsilon - Y_s^{\varepsilon'}, F_\varepsilon * \rho_s^\varepsilon(Y_s^\varepsilon) - F_{\varepsilon'} * \rho_s^{\varepsilon'}(Y_s^{\varepsilon'}) \rangle ds \\ &\leq 2 \int_0^t |Y_s^\varepsilon - Y_s^{\varepsilon'}| \left| F_\varepsilon * \rho_s^\varepsilon(Y_s^\varepsilon) - F_{\varepsilon'} * \rho_s^{\varepsilon'}(Y_s^{\varepsilon'}) \right| ds \\ &\leq 2 \int_0^t |Y_s^\varepsilon - Y_s^{\varepsilon'}| \left( C\omega(\varepsilon) + C\omega(\mathbb{P}\text{-ess sup } |Y_s^\varepsilon - Y_s^{\varepsilon'}|) \right) ds \\ &\leq C\omega(\varepsilon)T + C \int_0^t |Y_s^\varepsilon - Y_s^{\varepsilon'}| \omega(\mathbb{P}\text{-ess sup } |Y_s^\varepsilon - Y_s^{\varepsilon'}|) ds, \end{aligned}$$

where  $C$  only depends on  $D$ ,  $\|\rho_t^\varepsilon\|_\infty$ , and  $\|\rho_t^{\varepsilon'}\|_\infty$ . Hence, by Lemma 3.5, there exist constants  $C_T, \varepsilon_0$  depending only on  $C$  and  $T$  such that if  $\varepsilon < \varepsilon_0$ , then  $\omega(\varepsilon) < \Lambda_0$  and the following holds

$$\sup_{t \in [0, T]} \mathbb{P}\text{-ess sup } |Y_t^\varepsilon - Y_t^{\varepsilon'}| \leq C_T \omega(\varepsilon)^{\frac{1}{2}} e^{-CT}, \quad (4.25)$$

Therefore (4.24) holds. Hence there exists a limiting stochastic process  $Y_t \in C([0, T], D)$  such that

$$\sup_{t \in [0, T]} \mathbb{P}\text{-ess sup } |Y_t^\varepsilon - Y_t| \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0. \quad (4.26)$$

Assume that  $Y_t$  has the marginal density  $\rho_t$ . Then by the bound

$$\sup_{t \in [0, T]} \mathcal{W}_\infty(\rho_t^\varepsilon, \rho_t) \leq \sup_{t \in [0, T]} \mathbb{P}\text{-ess sup } |Y_t^\varepsilon - Y_t|,$$

we deduce that

$$\sup_{t \in [0, T]} \mathcal{W}_\infty(\rho_t^\varepsilon, \rho_t) \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0,$$

and  $\rho_t \in L^\infty(0, T; L^\infty(D))$ .

We are left to show that there exists  $\tilde{R}_t$  such that  $(Y_t, \tilde{R}_t)$  is a strong solution to the SDE (1.6). By Lemma 3.5 and (4.26),

$$\left| \int_0^t (F_\varepsilon * \rho_s^\varepsilon(Y_s^\varepsilon) - F * \rho_s(Y_s)) ds \right| \leq C \int_0^t \omega(\mathbb{P}\text{-ess sup } |Y_s^\varepsilon - Y_s|) ds + C\omega(\varepsilon)t \rightarrow 0 \quad \text{uniformly in } t \quad (4.27)$$

as  $\varepsilon \rightarrow 0$ . We define the boundary process  $\tilde{R}_t$  as

$$\tilde{R}_t := - \left( Y_t - Y_0 - \int_0^t F * \rho_s(Y_s) ds - \sqrt{2\nu} B_t \right), \quad (4.28)$$

such that  $(Y_t, \tilde{R}_t)$  is the strong solution to the SDE (1.6). The properties of  $R_t$  in (1.6) are verified in the same way as in [15, Step **B** on page 13], whose details are omitted here.  $\square$

Similar to Proposition 4.2, we show that any weak solution  $\rho_t \in L^\infty(0, T; L^\infty(D))$  to (1.1) can be represented as the family of time marginal density of a solution to (1.6).

**Proposition 4.3.** *If  $\rho_t \in L^\infty(0, T; L^\infty(D))$  is solution to (1.1) with the initial data  $\rho_0 \in L^\infty(D)$ , then given any random variable  $Y_0$  with the density  $\rho_0$ , there is a process  $Y_t$  with the marginal density  $\rho_t$  satisfying (1.6) with the initial data  $Y_0$ .*

*Proof.* Proposition 4.2 is a regularized version of this proposition, so the proof is almost the same. To this end, we change  $V_\varepsilon[g_t]$  into  $V[g_t] = \int_D F(x-y)g_t(y) dy$ . The details are omitted here.  $\square$

Using Proposition 4.3, we prove the stability estimates in Theorem 2.4.

*Proof of Theorem 2.4.* First we prove stability of the nonlinear SDE (1.6). Using similar arguments as in the proof of Theorem 2.3, we apply Itô's formula and get

$$\begin{aligned} |Y_t^1 - Y_t^2|^2 &= |Y_0^1 - Y_0^2|^2 + 2 \int_0^t \langle Y_s^1 - Y_s^2, F * \rho_s^1(Y_s^1) - F * \rho_s^2(Y_s^2) \rangle ds \\ &\quad - 2 \int_0^t \langle Y_s^1 - Y_s^2, n(Y_s^1) \rangle d|\tilde{R}^1|_s - 2 \int_0^t \langle Y_s^2 - Y_s^1, n(Y_s^2) \rangle d|\tilde{R}^2|_s \\ &\leq |Y_0^1 - Y_0^2|^2 + 2 \int_0^t \langle Y_s^1 - Y_s^2, F * \rho_s^1(Y_s^1) - F * \rho_s^2(Y_s^2) \rangle ds. \end{aligned} \quad (4.29)$$

Then it follows from Lemma 3.3 that there exists a constant  $C$  depending only on  $\|\rho^1\|_{L^\infty(0, T; L^\infty(D))}$  and  $\|\rho^2\|_{L^\infty(0, T; L^\infty(D))}$  such that

$$|Y_t^1 - Y_t^2|^2 \leq |Y_0^1 - Y_0^2|^2 + C \int_0^t |Y_s^1 - Y_s^2| \omega(\mathbb{P}\text{-ess sup } |Y_s^1 - Y_s^2|) ds. \quad (4.30)$$

The bound of  $|Y_t^1 - Y_t^2|$  is derived by considering three possible cases:

*Case 1:* There exists a small enough constant  $C_0(T)$  such that  $\mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2| \leq C_0(T)$ . Here  $C_0(T)$  is determined by Lemma 3.5. Actually, we can treat  $\mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2|$  as  $f_\varepsilon(t)$  in Lemma 3.5 and  $|Y_0^1 - Y_0^2|^2$  as  $C\omega(\varepsilon)T$ . Suppose that  $\mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2| < C_0(T) < 1$  sufficiently small. Then there exists two constants  $C_T$  and  $0 < \Lambda_0 < \frac{1}{2}$  depending only on  $C$  and  $T$  such that if  $|Y_0^1 - Y_0^2|^2 < C_0(T)^2 < \Lambda_0$ , then

$$\mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2| \leq C_T \{\mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2|\}^{\exp(-CT)} < 1 \quad \text{for all } t \in [0, T]. \quad (4.31)$$

*Case 2:* If  $\mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2| \geq C_0(T)$  and  $\mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2| \geq C_0(T)$  for any  $t \in [0, T]$ , then by the definition of  $\omega(r)$ , one obtains that there exists a constant  $C_1$  depending only on  $C_0(T)$  such that

$$\omega(\mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2|) \leq C_1 \mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2| \quad \text{for all } t \in [0, T]. \quad (4.32)$$

Thus it follows from (4.30) that

$$|Y_t^1 - Y_t^2|^2 \leq |Y_0^1 - Y_0^2|^2 + C_1 C \int_0^t |Y_s^1 - Y_s^2| \mathbb{P}\text{-ess sup } |Y_s^1 - Y_s^2| ds. \quad (4.33)$$

Applying the Gronwall inequality, one has for any  $t \in [0, T]$ ,

$$\mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2| \leq \mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2| (1 + CC_1 T e^{CC_1 T})^{\frac{1}{2}}. \quad (4.34)$$

*Case 3:* If  $\mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2| \geq C_0(T)$  and there exists a  $t_0 \in (0, T)$  such that  $\mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2| \geq C_0(T)$  for  $t \in [0, t_0)$  and  $\mathbb{P}\text{-ess sup } |Y_{t_0}^1 - Y_{t_0}^2| = C_0(T)$ . In this case, the estimate for the interval  $t \in [0, t_0)$  is reduced to the *Case 2* and one has

$$\mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2| \leq \mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2| (1 + CC_1 T e^{CC_1 T})^{\frac{1}{2}}. \quad (4.35)$$

Over the interval  $[t_0, T]$ , choosing  $t_0$  as a new initial time and repeating the proof in *Case 1* gives

$$\mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2| \leq C_T \{\mathbb{P}\text{-ess sup } |Y_{t_0}^1 - Y_{t_0}^2|\}^{\exp(-C(T-t_0))} \quad \text{for all } t \in [t_0, T]. \quad (4.36)$$

By (4.35) and the continuity of  $\mathbb{P}$ -ess sup  $|Y_t^1 - Y_t^2|$ , one has

$$\mathbb{P}\text{-ess sup } |Y_{t_0}^1 - Y_{t_0}^2| \leq \mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2| (1 + CC_1 T e^{CC_1 T})^{\frac{1}{2}}. \quad (4.37)$$

Combining (4.35), (4.36) and (4.37), we obtain

$$\mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2| \leq C_2 \max \left\{ \mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2|, \{ \mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2| \}^{\exp(-C_3 T)} \right\}, \quad (4.38)$$

for any  $t \in [0, T]$ . Combining (4.31), (4.34) and (4.38) finishes the proof of (2.7).

Next we prove the stability of the KS equation (1.1). Let  $\rho^1, \rho^2 \in L^\infty(0, T; L^\infty(D))$  be two weak solutions to (1.1) with the initial conditions  $\rho_0^1, \rho_0^2 \in L^\infty(D)$  respectively. Take two initial random variables  $Y_0^1$  and  $Y_0^2$  such that

$$\mathcal{W}_\infty(\rho_0^1, \rho_0^2) = \mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2|. \quad (4.39)$$

By Proposition 4.3, we build two self-consistent SDEs such that

$$\begin{cases} Y_t^i = Y_0^i + \int_0^t \int_D F(Y_s^i - y) \rho_s^i(y) dy ds + \sqrt{2\nu} B_t - \tilde{R}_t^i, & t > 0, i = 1, 2, \\ \tilde{R}_t^i = \int_0^t n(Y_s^i) d\tilde{R}^i|_s, & |\tilde{R}^i|_t = \int_0^t \mathbf{1}_{\partial D}(Y_s^i) d|\tilde{R}^i|_s, \end{cases}$$

which have unique strong solutions  $(Y_t^i)_{t \geq 0}$  with  $(\rho_t^i)_{t \geq 0}$  being the marginal density of  $(Y_t^i)_{t \geq 0}$ . Hence it follows from (2.7) and (4.39) that

$$\begin{aligned} \sup_{t \in [0, T]} \mathcal{W}_\infty(\rho_t^1, \rho_t^2) &\leq \sup_{t \in [0, T]} \mathbb{P}\text{-ess sup } |Y_t^1 - Y_t^2| \\ &\leq C_1 \max \left\{ \mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2|, \{ \mathbb{P}\text{-ess sup } |Y_0^1 - Y_0^2| \}^{\exp(-C_2 T)} \right\} \\ &= C_1 \max \left\{ \mathcal{W}_\infty(\rho_0^1, \rho_0^2), \{ \mathcal{W}_\infty(\rho_0^1, \rho_0^2) \}^{\exp(-C_2 T)} \right\}, \end{aligned}$$

which leads to the desired stability estimate (2.8).  $\square$

## 5. PROPAGATION OF CHAOS

In this section, we show an estimate of the collision time between particles and prove the propagation of chaos result stated in Theorem 2.7.

**5.1. Collision time between particles.** It is known that in  $\mathbb{R}^2$ , if  $8\pi\nu < \|\rho_0\|_1 = 1$ , then solutions to the Keller-Segel equation (1.1) concentrate within a finite time [18]. In this part, we show a similar bound for the expectation of the collision time for the interacting particle system (1.4).

**Theorem 5.1.** *Assume  $d = 2$  and  $\{X_0^i\}_{i=1}^N$  be  $N$  i.i.d. random variables with a common density  $\rho_0$ . Let  $\{(X_t^i, R_t^i)\}_{i=1}^N$  be weak solutions to (1.4) with the initial data  $\{X_0^i\}_{i=1}^N$ . For any fixed  $T > 0$  and  $8\pi\nu < 1$ , we define*

$$A(t) := \inf_{s \in [0, t]} \min_{i \neq j} |X_s^i - X_s^j|, \quad \tau_\varepsilon = \begin{cases} 0 & \text{if } \varepsilon \geq A(0), \\ \sup\{t \wedge T : A(t) \geq \varepsilon\} & \text{if } \varepsilon < A(0), \end{cases} \quad (5.1)$$

where  $t \wedge T = \min\{t, T\}$ . Let  $\tau = \lim_{\varepsilon \rightarrow 0} \tau_\varepsilon$ . Then the expectation of the collision time satisfies

$$\mathbb{E}(\tau) \leq \frac{2\pi\mathbb{E}[|X_0^1 - X_0|^2]}{1 - 8\pi\nu}, \quad X_0 := \frac{1}{N} \sum_{i=1}^N X_0^i. \quad (5.2)$$

*Proof.* Since system (1.4) has a weak solution until the collision time  $\tau$  and  $F_\varepsilon(x) = F(x)$  for any  $|x| > \varepsilon$ , we know that  $X_t^i \equiv X_t^{i, \varepsilon}$  and  $F_\varepsilon(x) = F(x)$  for  $t \in [0, \tau_\varepsilon]$ , where  $\{X_t^{i, \varepsilon}\}_{i=1}^N$  is a global solution to the regularized particle system (2.2). Let us denote

$$X_0 := \frac{1}{N} \sum_{i=1}^N X_0^i.$$



By the Itô's formula and (2.2), for  $t \in [0, \tau_\varepsilon]$ , one has

$$\begin{aligned}
|X_t^{i,\varepsilon} - X_0|^2 &= |X_0^i - X_0|^2 + 2 \int_0^t \left\langle X_s^{i,\varepsilon} - X_0, \frac{1}{N-1} \sum_{j \neq i}^N F(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) \right\rangle ds \\
&\quad + 2 \int_0^t \langle X_s^{i,\varepsilon} - X_0, \sqrt{2\nu} dB_s^{i,\varepsilon} \rangle - 2 \int_0^t \langle X_s^{i,\varepsilon} - X_0, dR_s^{i,\varepsilon} \rangle + 4\nu t \\
&\leq |X_0^i - X_0|^2 + 2 \int_0^t \left\langle X_s^{i,\varepsilon} - X_0, \frac{1}{N-1} \sum_{j \neq i}^N F(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) \right\rangle ds \\
&\quad + 2 \int_0^t \langle X_s^{i,\varepsilon} - X_0, \sqrt{2\nu} dB_s^{i,\varepsilon} \rangle + 4\nu t,
\end{aligned} \tag{5.3}$$

where we have used the convexity of the domain  $D$  such that

$$\int_0^t \langle X_s^{i,\varepsilon} - X_0, dR_s^i \rangle = \int_0^t \langle X_s^{i,\varepsilon} - X_0, n(X_t^{i,\varepsilon}) \rangle d|R^i|_s \geq 0. \tag{5.4}$$

Then summing (5.3) for all  $i$ 's, one has

$$\begin{aligned}
\sum_{i=1}^N |X_t^{i,\varepsilon} - X_0|^2 &\leq \sum_{i=1}^N |X_0^i - X_0|^2 + \frac{2}{N-1} \sum_{\substack{i,j=1 \\ i \neq j}}^N \int_0^t \langle X_s^{i,\varepsilon} - X_0, F(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) \rangle ds \\
&\quad + 2\sqrt{2\nu} \sum_{i=1}^N \int_0^t \langle X_s^{i,\varepsilon} - X_0, dB_s^{i,\varepsilon} \rangle + 4\nu Nt.
\end{aligned} \tag{5.5}$$

Next we verify that  $\int_0^{\tau_\varepsilon} \mathbb{E}[|X_t^{i,\varepsilon} - X_0|^2] dt < \infty$  which implies that [22, pp.28, Theorem 1]

$$\mathbb{E} \left[ \int_0^t \langle X_s^{i,\varepsilon} - X_0, dB_s^{i,\varepsilon} \rangle \right] = 0. \tag{5.6}$$

Indeed, notice that

$$\frac{2}{N-1} \sum_{\substack{i,j=1 \\ i \neq j}}^N \langle X_s^{i,\varepsilon} - X_0, F(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) \rangle = \frac{1}{N-1} \sum_{\substack{i,j=1 \\ i \neq j}}^N \langle X_s^{i,\varepsilon} - X_s^{j,\varepsilon}, F(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) \rangle = -\frac{N}{2\pi}. \tag{5.7}$$

By (5.5) and the exchangeability of  $\{(X_t^{i,\varepsilon})_{t \geq 0}\}_{i=1}^N$ , one has

$$\begin{aligned}
\mathbb{E}[|X_t^{i,\varepsilon} - X_0|^2] &\leq \mathbb{E}[|X_0^i - X_0|^2] + 4\nu t - \frac{t}{2\pi} + 2\sqrt{2\nu} \left\{ \mathbb{E} \left[ \left| \int_0^t \langle X_s^{i,\varepsilon} - X_0, dB_s^{i,\varepsilon} \rangle \right|^2 \right] \right\}^{\frac{1}{2}} \\
&\leq \mathbb{E}[|X_0^i - X_0|^2] + 4\nu t + 2\sqrt{2\nu} \left( \int_0^t \mathbb{E}[|X_s^{i,\varepsilon} - X_0|^2] ds \right)^{\frac{1}{2}},
\end{aligned}$$

with the last inequality coming from the Itô's isometry. Hence it follows from Gronwall inequality that

$$\int_0^{\tau_\varepsilon} \mathbb{E}[|X_t^{i,\varepsilon} - X_0|^2] dt < \infty. \tag{5.8}$$

Taking expectation of (5.5) and choosing  $t = \tau_\varepsilon$  in (5.5), one has

$$0 \leq \mathbb{E}[|X_{\tau_\varepsilon}^{i,\varepsilon} - X_0|^2] \leq \mathbb{E}[|X_0^i - X_0|^2] + \left(4\nu - \frac{1}{2\pi}\right) \mathbb{E}[\tau_\varepsilon] = \mathbb{E}[|X_0^1 - X_0|^2] + \left(4\nu - \frac{1}{2\pi}\right) \mathbb{E}[\tau_\varepsilon]. \tag{5.9}$$

By the positivity of left hand and  $8\pi\nu < 1$ , we obtain

$$\mathbb{E}[\tau_\varepsilon] \leq \frac{2\pi \mathbb{E}[|X_0^1 - X_0|^2]}{1 - 8\pi\nu}. \tag{5.10}$$

Hence, the desired bound for  $\tau$  holds by the monotone convergence theorem.  $\square$

In order to prove Theorem 2.7, we first show the following lemma:

**Lemma 5.1.** *Let  $\{X_t^{i,\varepsilon}\}_{i=1}^N$  be  $N$  exchangeable random variables on  $D$  satisfying (2.2). Let  $\{Y_t^{i,\varepsilon}\}_{i=1}^N$  be  $N$  i.i.d. random variables on  $D$  satisfying (4.19) with the common density  $\rho_t^\varepsilon \in L^\infty(0, T; L^\infty(D))$ . Then there exists a constant  $C > 0$  depending only on  $T, D$  and  $\|\rho^\varepsilon\|_{L^\infty(0, T; L^\infty(D))}$  such that*

$$\sup_{i=1, \dots, N} \left| \frac{1}{N-1} \sum_{j \neq i}^N F_\varepsilon(X_t^{i,\varepsilon} - X_t^{j,\varepsilon}) - \int_D F_\varepsilon(Y_t^{i,\varepsilon} - y) \rho_t^\varepsilon(y) dy \right| \leq C\varepsilon^{-d} \sup_{i=1, \dots, N} |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}| + \mathcal{H}_N,$$

where  $\mathcal{H}_N$  is given by

$$\mathcal{H}_N := \sup_{i=1, \dots, N} \left| \frac{1}{N-1} \sum_{j \neq i}^N F_\varepsilon(Y_t^{i,\varepsilon} - Y_t^{j,\varepsilon}) - \int_D F_\varepsilon(Y_t^{i,\varepsilon} - y) \rho_t^\varepsilon(y) dy \right|, \quad (5.11)$$

and it satisfies

$$\mathbb{E}[|\mathcal{H}_N|^2] \leq \frac{C\varepsilon^{-2(d-1)}}{N-1}. \quad (5.12)$$

*Proof.* For  $i = 1$ , we split the error

$$\begin{aligned} & \left| \frac{1}{N-1} \sum_{j=2}^N F_\varepsilon(X_t^{1,\varepsilon} - X_t^{j,\varepsilon}) - \int_D F_\varepsilon(Y_t^{1,\varepsilon} - y) \rho_t^\varepsilon(y) dy \right| \\ & \leq \left| \frac{1}{N-1} \sum_{j=2}^N F_\varepsilon(X_t^{1,\varepsilon} - X_t^{j,\varepsilon}) - \frac{1}{N-1} \sum_{j=2}^N F_\varepsilon(Y_t^{1,\varepsilon} - Y_t^{j,\varepsilon}) \right| \\ & \quad + \left| \frac{1}{N-1} \sum_{j=2}^N F_\varepsilon(Y_t^{1,\varepsilon} - Y_t^{j,\varepsilon}) - \int_D F_\varepsilon(Y_t^{1,\varepsilon} - y) \rho_t^\varepsilon(y) dy \right| =: |I_1| + |I_2|. \end{aligned} \quad (5.13)$$

By Lemma 3.1 (b) one has

$$|I_1| \leq \frac{1}{N-1} \sum_{j=2}^N C\varepsilon^{-d} (|X_t^{1,\varepsilon} - Y_t^{1,\varepsilon}| + |X_t^{j,\varepsilon} - Y_t^{j,\varepsilon}|) = C \frac{\varepsilon^{-d}}{N-1} \sum_{j=1}^N |X_t^{j,\varepsilon} - Y_t^{j,\varepsilon}|. \quad (5.14)$$

This shows

$$|I_1| \leq C\varepsilon^{-d} \sup_{i=1, \dots, N} |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}|. \quad (5.15)$$

Next, we rewrite  $I_2$  as

$$I_2 = \frac{1}{N-1} \sum_{j=2}^N Z_j, \quad (5.16)$$

with

$$Z_j := F_\varepsilon(Y_t^{1,\varepsilon} - Y_t^{j,\varepsilon}) - \int_D F_\varepsilon(Y_t^{1,\varepsilon} - y) \rho_t^\varepsilon(y) dy, \quad j \geq 2. \quad (5.17)$$

Since  $\{Y_t^{j,\varepsilon}\}_{j=1}^N$  are i.i.d. random variables, it holds that

$$\mathbb{E}[Z_j Z_k] = 0, \quad j \neq k. \quad (5.18)$$

To see this, we consider  $Y_t^{1,\varepsilon}$  is given and denote

$$\mathbb{E}'[\cdot] = \mathbb{E}[\cdot | Y_t^{1,\varepsilon}]. \quad (5.19)$$

Since  $Z_j$  and  $Z_k$  are independent for  $j \neq k$ , it holds that  $\mathbb{E}'[Z_j Z_k] = \mathbb{E}'[Z_j] \mathbb{E}'[Z_k] = 0$  when  $j \neq k$ . Hence

$$\mathbb{E}[Z_j Z_k] = \mathbb{E}_1 \mathbb{E}'[Z_j Z_k] = 0, \quad j \neq k, \quad (5.20)$$

where  $\mathbb{E}_1$  means taking expectation on  $Y_t^{1,\varepsilon}$ .

Hence, one concludes that

$$\mathbb{E}[|I_2|^2] \leq \frac{1}{N-1} \mathbb{E}[|Z_2|^2] \leq \frac{2}{N-1} \mathbb{E} \left[ F_\varepsilon^2(Y_t^{1,\varepsilon} - Y_t^{2,\varepsilon}) + \left( \int_D F_\varepsilon(Y_t^{1,\varepsilon} - y) \rho_t^\varepsilon(y) dy \right)^2 \right] \leq \frac{C\varepsilon^{-2(d-1)}}{N-1}, \quad (5.21)$$

where we have used Lemma 3.1 (b) and Lemma 3.2 (a). Combining estimates (5.15) and (5.21), we have

$$\left| \frac{1}{N-1} \sum_{j=2}^N F_\varepsilon(X_t^{1,\varepsilon} - X_t^{j,\varepsilon}) - \int_D F_\varepsilon(Y_t^{1,\varepsilon} - y) \rho_t^\varepsilon(y) dy \right| \leq C\varepsilon^{-d} \sup_{i=1, \dots, N} |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}| + |I_2|, \quad (5.22)$$

with the bound (5.21) for  $I_2$ . The same estimate holds for the case  $i = 2, \dots, N$ .  $\square$

To prove the propagation of chaos, we need to introduce an auxiliary stochastic mean-field dynamics  $\{Y_t^i\}_{i=1}^N$  satisfying

$$\begin{cases} Y_t^i = Y_0^i + \int_0^t \int_D F(Y_s^i - y) \rho_s(y) dy ds + \sqrt{2\nu} B_t^i - \tilde{R}_t^i, & t > 0, \quad i = 1, \dots, N, \\ \tilde{R}_t^i = \int_0^t n(Y_s^i) d|\tilde{R}^i|_s, \quad |\tilde{R}^i|_t = \int_0^t \mathbf{1}_{\partial D}(Y_s^i) d|\tilde{R}^i|_s, \end{cases} \quad (5.23)$$

where  $(\rho_t)_{t \geq 0}$  is the common marginal density of  $(\{Y_t^i\}_{i=1}^N)_{t \geq 0}$  for any  $t \geq 0$ . Here we set the initial date  $(Y_0^i)_{i=1, \dots, N} = (X_0^i)_{i=1, \dots, N}$  i.i.d. sharing the same density  $\rho_0(x)$ . In other words here  $(\{Y_t^i\}_{i=1}^N)_{t \geq 0}$  are  $N$  copies of the strong solutions to (1.6).

Finally, we present the proof of Theorem 2.7 regarding the propagation of chaos.

*Proof of Theorem 2.7.* It follows from Theorem 2.1, Theorem 2.3 and Theorem 2.4 that there exists a weak solution  $\{X_t^{i,\varepsilon}\}_{i=1}^N$  to (2.2) and a unique pathwise solution  $\{Y_t^i\}_{i=1}^N$  to (5.23) with marginal density  $\rho_t$  on the time interval  $[0, T]$  for some  $T > 0$ , such that  $\rho_t \in L^\infty(0, T; L^\infty(D))$ . This implies that we are able to define solutions for those two equations on the same probability space with the same initial condition and Brownian motion.

Let  $\{Y_t^{i,\varepsilon}\}_{i=1}^N$  be the  $N$  copies of solutions to regularized mean-field dynamics (4.19). Then it follows from the same argument as in the proof of (4.25), we obtain that there exists constants  $C_T, \varepsilon_0$  depending only on  $T, \|\rho^\varepsilon\|_{L^\infty(0, T; L^\infty(D))}$  and  $\|\rho\|_{L^\infty(0, T; L^\infty(D))}$  such that if  $\varepsilon < \varepsilon_0$ , then for any  $1 \leq i \leq N$

$$\sup_{t \in [0, T]} \mathbb{P}\text{-ess sup} |Y_t^{i,\varepsilon} - Y_t^i| \leq C_T \omega(\varepsilon)^{\frac{1}{2}} e^{-C_T}. \quad (5.24)$$

Using the similar argument as in the proof of Theorem 2.3, it follows from Itô's formula that

$$\begin{aligned} |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}|^2 &= 2 \int_0^t \left\langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, \frac{1}{N-1} \sum_{j \neq i}^N F_\varepsilon(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) - \int_D F_\varepsilon(Y_s^{i,\varepsilon} - y) \rho_t^\varepsilon(y) dy \right\rangle ds \\ &\quad - 2 \int_0^t \langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, n(X_s^{i,\varepsilon}) \rangle d|R^{i,\varepsilon}|_s - 2 \int_0^t \langle Y_s^{i,\varepsilon} - X_s^{i,\varepsilon}, n(Y_s^{i,\varepsilon}) \rangle d|\tilde{R}^{i,\varepsilon}|_s \\ &\quad + 2 \int_0^t \langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, dB_s^{i,\varepsilon} \rangle - 2 \int_0^t \langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, dB_s^i \rangle \\ &\leq 2 \int_0^t \left\langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, \frac{1}{N-1} \sum_{j \neq i}^N F_\varepsilon(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) - \int_D F_\varepsilon(Y_s^{i,\varepsilon} - y) \rho_t^\varepsilon(y) dy \right\rangle ds \\ &\quad + 2 \int_0^t \langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, dB_s^{i,\varepsilon} \rangle - 2 \int_0^t \langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, dB_s^i \rangle, \end{aligned} \quad (5.25)$$

where in the second inequality we have used the convexity of the domain  $D$ . Applying Lemma 5.1, one finds

$$\begin{aligned}
|X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}|^2 &\leq 2 \int_0^t |X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}| \left( C\varepsilon^{-d} \sup_{i=1,\dots,N} |X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}| + \mathcal{H}_N \right) ds \\
&\quad + 2 \int_0^t \langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, dB_s^{i,\varepsilon} \rangle - 2 \int_0^t \langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, dB_s^i \rangle \\
&\leq C\varepsilon^{-d} \int_0^t \sup_{i=1,\dots,N} |X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}|^2 ds + C\mathcal{H}_N T \\
&\quad + 2 \int_0^t \langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, dB_s^{i,\varepsilon} \rangle - 2 \int_0^t \langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, dB_s^i \rangle.
\end{aligned} \tag{5.26}$$

If we can show that

$$\mathbb{E} \left[ |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}|^2 \right] < +\infty, \tag{5.27}$$

then according to [22, pp.28, Theorem 1] one has

$$\mathbb{E} \left[ \int_0^t \langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, dB_s^{i,\varepsilon} \rangle \right] = \mathbb{E} \left[ \int_0^t \langle X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}, dB_s^i \rangle \right] = 0. \tag{5.28}$$

In order to prove (5.27), take the expectation on (5.26). This gives

$$\mathbb{E} \left[ \sup_{i=1,\dots,N} |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}|^2 \right] \leq C\varepsilon^{-d} \int_0^t \mathbb{E} \left[ \sup_{i=1,\dots,N} |X_s^{i,\varepsilon} - Y_s^{i,\varepsilon}|^2 \right] ds + \frac{CT\varepsilon^{-(d-1)}}{\sqrt{N-1}}, \tag{5.29}$$

where we have used

$$\mathbb{E}[|\mathcal{H}_N|] \leq (\mathbb{E}[|\mathcal{H}_N|^2])^{\frac{1}{2}} \leq \frac{C\varepsilon^{-(d-1)}}{\sqrt{N-1}}. \tag{5.30}$$

Applying Gronwall's inequality, we have

$$\mathbb{E} \left[ \sup_{i=1,\dots,N} |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}| \right] \leq \left( \mathbb{E} \left[ \sup_{i=1,\dots,N} |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}|^2 \right] \right)^{\frac{1}{2}} \leq \frac{CT\varepsilon^{-d+\frac{1}{2}} \exp(CT\varepsilon^{-d})}{(N-1)^{\frac{1}{4}}} \tag{5.31}$$

for all  $t \in [0, T]$ . Let  $\varepsilon = (\frac{\log N}{8CT})^{-\frac{1}{d}}$ , then one has

$$\mathbb{E} \left[ \sup_{i=1,\dots,N} |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}| \right] \leq \left( \mathbb{E} \left[ \sup_{i=1,\dots,N} |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}|^2 \right] \right)^{\frac{1}{2}} \leq \frac{N^{\frac{1}{8}} \log(N)}{8CT(N-1)^{\frac{1}{4}}} \leq CN^{-\frac{1}{8}} \log(N), \tag{5.32}$$

which verifies (5.27) thus (5.28).

Now recall the definition of the empirical measures

$$\mu_t^{X,\varepsilon} = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,\varepsilon}}, \quad \mu_t^{Y,\varepsilon} = \frac{1}{N} \sum_{i=1}^N \delta_{Y_t^{i,\varepsilon}}, \quad \mu_t^Y = \frac{1}{N} \sum_{i=1}^N \delta_{Y_t^i}, \tag{5.33}$$

and notice that

$$\pi_* = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,\varepsilon}} \delta_{Y_t^{i,\varepsilon}} \in \Lambda(\mu_t^{X,\varepsilon}, \mu_t^{Y,\varepsilon}). \tag{5.34}$$

Then by the definition of  $\mathcal{W}_\infty$  distance one has

$$\mathcal{W}_\infty(\mu_t^{X,\varepsilon} - \mu_t^{Y,\varepsilon}) \leq \pi_* \text{-ess sup}_{(x,y) \in D \times D} |x - y| \leq \sup_{i=1,\dots,N} |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}|. \tag{5.35}$$

This leads to

$$\mathbb{E} \left[ \mathcal{W}_\infty(\mu_t^{X,\varepsilon} - \mu_t^{Y,\varepsilon}) \right] \leq \mathbb{E} \left[ \sup_{i=1,\dots,N} |X_t^{i,\varepsilon} - Y_t^{i,\varepsilon}| \right] \leq CN^{-\frac{1}{8}} \log(N) \tag{5.36}$$

and

$$\mathbb{E} \left[ \mathcal{W}_\infty(\mu_t^{Y,\varepsilon} - \mu_t^Y) \right] \leq \mathbb{E} \left[ \sup_{i=1,\dots,N} |Y_t^{i,\varepsilon} - Y_t^i| \right] \leq CT\omega(\varepsilon)e^{-CT} \leq C[\omega(\log^{-\frac{1}{d}}(N))]^{\exp(-CT)}. \tag{5.37}$$

Furthermore, it follows from the convergence estimate obtained in [20, Theorem 1] that

$$\mathbb{E} [\mathcal{W}_2^2(\mu_t^Y - \rho_t)] \leq C \begin{cases} N^{-\frac{1}{2}} + N^{-\frac{q-2}{q}}, & \text{if } d < 4 \text{ and } q \neq 4, \\ N^{-\frac{1}{2}} \log(N+1) + N^{-\frac{q-2}{q}}, & \text{if } d = 4 \text{ and } q \neq 4, \\ N^{-\frac{2}{d}} + N^{-\frac{q-2}{q}}, & \text{if } d > 4 \text{ and } q \neq \frac{d}{d-2}. \end{cases} \quad (5.38)$$

Combining inequalities (5.36) and (5.37), we arrive at

$$\begin{aligned} & \mathbb{E} [\mathcal{W}_2(\mu_t^{X,\varepsilon} - \rho_t)] \leq \mathbb{E} [\mathcal{W}_2(\mu_t^{X,\varepsilon} - \mu_t^{Y,\varepsilon})] + \mathbb{E} [\mathcal{W}_2(\mu_t^{Y,\varepsilon} - \mu_t^Y)] + \mathbb{E} [\mathcal{W}_2(\mu_t^Y - \rho_t)] \\ & \leq C \mathbb{E} [\mathcal{W}_\infty(\mu_t^{X,\varepsilon} - \mu_t^{Y,\varepsilon})] + C \mathbb{E} [\mathcal{W}_\infty(\mu_t^{Y,\varepsilon} - \mu_t^Y)] + \mathbb{E} [\mathcal{W}_2(\mu_t^Y - \rho_t)] \\ & \leq CN^{-\frac{1}{8}}(\log N) + C[\omega(\log^{-\frac{1}{d}}(N))]^{\exp(-CT)} \\ & \quad + C \begin{cases} N^{-\frac{1}{4}} + N^{-\frac{q-2}{2q}}, & \text{if } d < 4 \text{ and } q \neq 4, \\ N^{-\frac{1}{4}}(\log(N+1))^{\frac{1}{2}} + N^{-\frac{q-2}{2q}}, & \text{if } d = 4 \text{ and } q \neq 4, \\ N^{-\frac{1}{d}} + N^{-\frac{q-2}{2q}}, & \text{if } d > 4 \text{ and } q \neq \frac{d}{d-2}, \end{cases} \\ & \leq C[\omega(\log^{-\frac{1}{d}}(N))]^{\exp(-CT)}, \end{aligned} \quad (5.39)$$

which completes the proof.  $\square$

#### REFERENCES

- [1] Jacob Bedrossian, Nancy Rodríguez, and Andrea L Bertozzi, *Local and global well-posedness for aggregation equations and Patlak-Keller-Segel models with degenerate diffusion*, *Nonlinearity* **24** (2011), no. 6, 1683.
- [2] Andrea L Bertozzi and Dejan Slepčev, *Existence and uniqueness of solutions to an aggregation equation with degenerate diffusion*, *Communications on Pure and Applied Analysis* **9** (2009), no. 6, 1617.
- [3] Shen Bian and Jian-Guo Liu, *Dynamic and steady states for multi-dimensional Keller-Segel model with diffusion exponent  $m > 0$* , *Communications in Mathematical Physics* **323** (2013), no. 3, 1017–1070.
- [4] Piotr Biler, *Existence and nonexistence of solutions for a model of gravitational interaction of particles, III*, *Colloquium mathematicae*, 1995, pp. 229–239.
- [5] Piotr Biler, *Local and global solvability of some parabolic systems modelling chemotaxis*, *Advances in Mathematical Sciences and Applications* **8** (1998), 715–743.
- [6] Adrien Blanchet, José A Carrillo, and Nader Masmoudi, *Infinite time aggregation for the critical Patlak-Keller-Segel model in  $\mathbb{R}^2$* , *Communications on Pure and Applied Mathematics* **61** (2008), no. 10, 1449–1481.
- [7] Niklas Boers and Peter Pickl, *On mean field limits for dynamical systems*, *Journal of Statistical Physics* **164** (2016), no. 1, 1–16.
- [8] François Bolley, José A Canizo, and José A Carrillo, *Stochastic mean-field limit: non-Lipschitz forces and swarming*, *Mathematical Models and Methods in Applied Sciences* **21** (2011), no. 11, 2179–2210.
- [9] W Braun and K Hepp, *The Vlasov dynamics and its fluctuations in the  $\frac{1}{N}$  limit of interacting classical particles*, *Communications in mathematical physics* **56** (1977), no. 2, 101–113.
- [10] José A Carrillo, Young-Pil Choi, and Samir Salem, *Propagation of chaos for the VFPF equation with a polynomial cut-off*, arXiv preprint arXiv:1802.01929 (2018).
- [11] José A Carrillo, Massimo Fornasier, Giuseppe Toscani, and Francesco Vecil, *Particle, kinetic, and hydrodynamic models of swarming*, *Mathematical modeling of collective behavior in socio-economic and life sciences* (2010), 297–336.
- [12] José A Carrillo, Robert J McCann, Cédric Villani, et al., *Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates*, *Revista Matemática Iberoamericana* **19** (2003), no. 3, 971–1018.
- [13] José A Carrillo, Dejan Slepčev, and Lijiang Wu, *Nonlocal-interaction equations on uniformly prox-regular sets*, *Discrete and Continuous Dynamical Systems-Series A* **36** (2014), no. 3, 1209–1247.
- [14] José Antonio Carrillo, Stefano Lisini, and Edoardo Mainini, *Uniqueness for Keller-Segel-type chemotaxis models*, *Dynamical Systems* **34** (2014), no. 4, 1319–1338.

- [15] Young-Pil Choi and Samir Salem, *Propagation of chaos for aggregation equations with no-flux boundary conditions and sharp sensing zones*, Mathematical Models and Methods in Applied Sciences **28** (2018), no. 02, 223–258.
- [16] Tomasz Cieślak and Michael Winkler, *Global bounded solutions in a two-dimensional quasilinear Keller-Segel system with exponentially decaying diffusivity and subcritical sensitivity*, Nonlinear Analysis: Real World Applications **35** (2017), 1–19.
- [17] Roland L'vovich Dobrushin, *Vlasov equations*, Functional Analysis and Its Applications **13** (1979), no. 2, 115–123.
- [18] Jean Dolbeault and Benoît Perthame, *Optimal critical mass in the two dimensional Keller-Segel model in  $\mathbb{R}^2$* , Comptes Rendus Mathématique **339** (2004), no. 9, 611–616.
- [19] Giani Egaña Fernández and Stéphane Mischler, *Uniqueness and long time asymptotic for the Keller-Segel equation: the parabolic-elliptic case*, Archive for Rational Mechanics and Analysis **220** (2016), no. 3, 1159–1194.
- [20] Nicolas Fournier and Arnaud Guillin, *On the rate of convergence in Wasserstein distance of the empirical measure*, Probability Theory and Related Fields **162** (2015), no. 3-4, 707–738.
- [21] Ana García-Cañizares and Peter Pickl, *Microscopic derivation of the Keller-Segel equation in the sub-critical regime*, arXiv preprint arXiv:1703.04376 (2017).
- [22] Iosif Il'ich Gihman and Anatolií Vladimirovich Skorohod, *Stochastic differential equations*, Springer-Verlag Berlin Heidelberg, New York, 1972.
- [23] Dirk Horstmann et al., *From 1970 until present: the Keller-Segel model in chemotaxis and its consequences* (2003).
- [24] Hui Huang and Jian-Guo Liu, *Well-posedness for the Keller-Segel equation with fractional Laplacian and the theory of propagation of chaos*, Kinetic And Related Models **9** (2016), no. 4, 715–748.
- [25] Hui Huang and Jian-Guo Liu, *Discrete-in-time random particle blob method for the Keller-Segel equation and convergence analysis*, Communication in Mathematical Sciences **15** (2017), no. 7, 1821–1842.
- [26] Hui Huang and Jian-Guo Liu, *Error estimate of a random particle blob method for the Keller-Segel equation*, Mathematics of Computation **86** (2017), no. 308, 2719–2744.
- [27] Hui Huang, Jian-Guo Liu, and Jianfeng Lu, *Learning interacting particle systems: diffusion parameter estimation for aggregation equations*, arXiv preprint arXiv:1802.02267 (2018).
- [28] Pierre-Emmanuel Jabin, *A review of the mean field limits for Vlasov equations*, Kinet. Relat. Models **7** (2014), no. 4, 661–711.
- [29] Pierre-Emmanuel Jabin and Maxime Hauray, *Particles approximations of Vlasov equations with singular forces: Propagation of chaos.*, Annales scientifiques de l'école normale supérieure, 2015.
- [30] Pierre-Emmanuel Jabin and Zhenfu Wang, *Mean field limit for stochastic particle systems*, Active particles, volume 1, 2017, pp. 379–402.
- [31] Willi Jäger and Stephan Luckhaus, *On explosions of solutions to a system of partial differential equations modelling chemotaxis*, Transactions of the american mathematical society **329** (1992), no. 2, 819–824.
- [32] Mark Kac, *Foundations of kinetic theory*, Proceedings of the third berkeley symposium on mathematical statistics and probability, 1956, pp. 171–197.
- [33] Tosio Kato, *On classical solutions of the two-dimensional non-stationary euler equation*, Archive for Rational Mechanics and Analysis **25** (1967), no. 3, 188–200.
- [34] Tatsuki Kawakami and Yoshie Sugiyama, *Uniqueness theorem on weak solutions to the Keller-Segel system of degenerate and singular types*, Journal of Differential Equations **260** (2016), no. 5, 4683–4716.
- [35] Evelyn F Keller and Lee A Segel, *Initiation of slime mold aggregation viewed as an instability*, Journal of Theoretical Biology **26** (1970), no. 3, 399–415.
- [36] Dustin Lazarovici and Peter Pickl, *A mean field limit for the Vlasov-Poisson system*, Archive for Rational Mechanics and Analysis **225** (2017), 1201–1231.
- [37] Pierre-Louis Lions and Alain-Sol Sznitman, *Stochastic differential equations with reflecting boundary conditions*, Communications on Pure and Applied Mathematics **37** (1984), no. 4, 511–537.
- [38] Jian-Guo Liu and Jinhuan Wang, *Refined hyper-contractivity and uniqueness for the Keller-Segel equations*, Applied Mathematics Letters **52** (2016), 212–219.
- [39] Toshitaka Nagai, *Blowup of nonradial solutions to parabolic-elliptic systems modeling chemotaxis in two-dimensional domains*, Journal of Inequalities and Applications **2001** (2001), no. 1, 970292.
- [40] Bernt Øksendal, *Stochastic differential equations: An introduction with applications, sixth edition*, Springer, Berlin, 2003.
- [41] Benoît Perthame, *Transport equations in biology*, Springer Science & Business Media, 2006.
- [42] Yasumasa Saisho, *Stochastic differential equations for multi-dimensional domain with reflecting boundary*, Probability Theory and Related Fields **74** (1987), no. 3, 455–477.
- [43] Anatoliy V Skorokhod, *Stochastic equations for diffusion processes in a bounded region*, Theory of Probability & Its Applications **6** (1961), no. 3, 264–274.

- [44] Anatoliy V Skorokhod, *Stochastic equations for diffusion processes in a bounded region. II*, Theory of Probability & Its Applications **7** (1962), no. 1, 3–23.
- [45] Philippe Souplet and Michael Winkler, *Blow-up profiles for the parabolic-elliptic Keller-Segel system in dimensions  $n \geq 3$* , arXiv preprint arXiv:1802.03093 (2018).
- [46] Herbert Spohn, *Dynamics of charged particles and their radiation field*, Cambridge University Press, 2004.
- [47] Yoshie Sugiyama, *Uniqueness and regularity of weak solutions for the 1-d degenerate Keller-Segel systems*, Nonlinear Analysis: Theory, Methods & Applications **73** (2010), no. 7, 2230–2244.
- [48] Alain-Sol Sznitman, *Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated*, Journal of functional analysis **56** (1984), no. 3, 311–336.
- [49] Alain-Sol Sznitman, *Topics in propagation of chaos*, Ecole d'Eté de Probabilités de Saint-Flour XIX1989 (1991), 165–251.
- [50] Hiroshi Tanaka, *Stochastic differential equations with reflecting*, Stochastic Processes: Selected Papers of Hiroshi Tanaka **9** (1979), 157.
- [51] Youshan Tao, *Boundedness in a chemotaxis model with oxygen consumption by bacteria*, Journal of Mathematical Analysis and Applications **381** (2011), no. 2, 521–529.
- [52] Youshan Tao and Michael Winkler, *Eventual smoothness and stabilization of large-data solutions in a three-dimensional chemotaxis system with consumption of chemoattractant*, Journal of Differential Equations **252** (2012), no. 3, 2520–2543.
- [53] Cédric Villani, *Optimal transport: old and new*, Vol. 338, Springer Science & Business Media, 2008.
- [54] Michael Winkler, *The two-dimensional Keller-Segel system with singular sensitivity and signal absorption: Global large-data solutions and their relaxation properties*, Mathematical Models and Methods in Applied Sciences **26** (2016), no. 05, 987–1024.
- [55] Rong Yang and Jian-Guo Liu, *Propagation of chaos for the Keller-Segel equation with a logarithmic cut-off*. preprint.

DEPARTMENT OF MATHEMATICS, SIMON FRASER UNIVERSITY, 8888 UNIVERSITY DR., BURNABY, BC V5A 1S6, CANADA  
*E-mail address:* `van@math.sfu.ca`

DEPARTMENT OF MATHEMATICS, SIMON FRASER UNIVERSITY, 8888 UNIVERSITY DR., BURNABY, BC V5A 1S6, CANADA  
*E-mail address:* `hha101@sfu.ca`

DEPARTMENT OF MATHEMATICS, SIMON FRASER UNIVERSITY, 8888 UNIVERSITY DR., BURNABY, BC V5A 1S6, CANADA  
*E-mail address:* `weirans@sfu.ca`